

## **Proximal envelopes: Smooth optimization algorithms for nonsmooth problems**

**Lorenzo Stella**

Supervisors:

Prof. Panagiotis Patrinos (KU Leuven)

Prof. Alberto Bemporad (IMT Lucca)

Dissertation presented in partial fulfillment of the requirements for:

PhD in Computer, Decision and Systems Science, XXVIII Cycle

Degree of Doctor of Engineering Science (PhD): Electrical Engineering

**2017**



**The dissertation of Lorenzo Stella is approved.**

Program Coordinator: Prof. Pietro Pietrini, IMT Lucca

Supervisor: Prof. Panagiotis Patrinos, KU Leuven

Supervisor: Prof. Alberto Bemporad, IMT Lucca

The dissertation of Lorenzo Stella has been reviewed by:

Prof. Pontus Giselsson, Lund University

Prof. Johan Suykens, KU Leuven

**IMT School for Advanced Studies, Lucca**

**2017**



A Rebecca e Ruben



# Contents

|  |              |
|--|--------------|
| List of Figures  | x <i>i</i>   |
| List of Tables   | x <i>ii</i>  |
| Acknowledgements   | x <i>iii</i> |
| Vita and Publications  | x <i>iv</i>  |
| Abstract   | x <i>vi</i>  |
| Notation   | xv <i>ii</i> |
| <b>1 Introduction</b>  | <b>1</b>     |
| 1.1 Structured optimization . . . . .                          | 2            |
| 1.2 Proximal algorithms . . . . .                              | 5            |
| 1.3 Proximal envelopes . . . . .                               | 8            |
| 1.4 Contributions and organization . . . . .                   | 11           |
| 1.A Tools and notation . . . . .                               | 13           |
| 1.A.1 Convex and variational analysis tools . . . . .          | 13           |
| 1.A.2 Continuity and (generalized) differentiability . . . . . | 16           |
| 1.A.3 Convergence rates . . . . .                              | 18           |
| <b>2 Forward-backward quasi-Newton methods</b>                 | <b>20</b>    |
| 2.1 Introduction . . . . .                                     | 20           |

|          |   |           |
|----------|---|-----------|
| 2.1.1    | Forward-backward splitting . . . . .                        | 23        |
| 2.2      | Forward-backward envelope . . . . .                         | 26        |
| 2.2.1    | Basic inequalities . . . . .                                | 26        |
| 2.2.2    | First- and second-order properties . . . . .                | 31        |
| 2.2.3    | Interpretations . . . . .                                   | 36        |
| 2.3      | Forward-backward line-search methods . . . . .              | 37        |
| 2.3.1    | Convergence in the convex case . . . . .                    | 44        |
| 2.3.2    | Convergence under KL assumption . . . . .                   | 47        |
| 2.4      | Quasi-Newton methods . . . . .                              | 52        |
| 2.4.1    | BFGS . . . . .  | 57        |
| 2.4.2    | L-BFGS . . . . .  | 59        |
| 2.5      | Simulations . . . . .                                       | 60        |
| 2.5.1    | Lasso . . . . .   | 60        |
| 2.5.2    | Sparse logistic regression . . . . .                        | 62        |
| 2.5.3    | Group lasso . . . . .                                       | 64        |
| 2.5.4    | Matrix completion . . . . .                                 | 66        |
| 2.5.5    | Image restoration . . . . .                                 | 68        |
| 2.6      | Conclusions . . . . .                                       | 69        |
| 2.A      | Additional results . . . . .                                | 71        |
| <b>3</b> | <b>A simple and efficient algorithm for nonlinear MPC</b>   | <b>76</b> |
| 3.1      | Introduction . . . . .                                      | 76        |
| 3.1.1    | Problems framework and motivation . . . . .                 | 77        |
| 3.1.2    | Contributions . . . . .                                     | 78        |
| 3.2      | Newton-type forward-backward method . . . . .               | 79        |
| 3.2.1    | Newton-type methods on generalized equations . . . . .      | 80        |
| 3.2.2    | Forward-backward envelope . . . . .                         | 81        |
| 3.2.3    | A superlinearly convergent algorithm based on FBS . . . . . | 83        |
| 3.3      | Nonlinear MPC . . . . .                                     | 88        |
| 3.3.1    | Handling state constraints . . . . .                        | 89        |
| 3.4      | Numerical Simulations . . . . .                             | 92        |
| 3.4.1    | Simulation scenario . . . . .                               | 93        |



|          |  |            |
|----------|--|------------|
| 3.4.2    | Results . . . . .  | 95         |
| 3.5      | Conclusions . . . . .                                    | 96         |
| <b>4</b> | <b>Newton-type alternating minimization algorithm</b>    | <b>99</b>  |
| 4.1      | Introduction . . . . .                                   | 99         |
| 4.1.1    | Contributions . . . . .                                  | 101        |
| 4.2      | Background and proposed algorithm . . . . .              | 102        |
| 4.2.1    | Newton-type alternating minimization algorithm . . . . . | 105        |
| 4.2.2    | Quasi-Newton directions . . . . .                        | 107        |
| 4.3      | Alternating minimization envelope . . . . .              | 109        |
| 4.3.1    | Analogy with the dual Moreau envelope . . . . .          | 111        |
| 4.4      | Convergence . . . . .                                    | 112        |
| 4.5      | First- and second-order properties . . . . .             | 118        |
| 4.6      | Superlinear convergence . . . . .                        | 124        |
| 4.7      | Simulations . . . . .                                    | 127        |
| 4.7.1    | Linear MPC . . . . .                                     | 128        |
| 4.8      | Conclusions . . . . .                                    | 133        |
| 4.A      | Additional results . . . . .                             | 135        |
| <b>5</b> | <b>Fast Douglas-Rachford splitting algorithm</b>         | <b>139</b> |
| 5.1      | Introduction . . . . .                                   | 139        |
| 5.1.1    | Contributions . . . . .                                  | 141        |
| 5.2      | Douglas-Rachford envelope . . . . .                      | 142        |
| 5.2.1    | DRS as a variable-metric gradient method . . . . .       | 145        |
| 5.2.2    | Connection between DRS and FBS . . . . .                 | 146        |
| 5.3      | Convergence rate and stepsize selection in DRS . . . . . | 148        |
| 5.4      | Fast Douglas-Rachford splitting . . . . .                | 152        |
| 5.5      | Simulations . . . . .                                    | 154        |
| 5.5.1    | Box-constrained QP . . . . .                             | 154        |
| 5.5.2    | Lasso . . . . .  | 155        |
| 5.6      | Conclusions and future work . . . . .                    | 156        |
| 5.A      | Proofs and additional results . . . . .                  | 157        |

|          |                                |            |
|----------|--------------------------------|------------|
| <b>6</b> | <b>Conclusions and outlook</b> | <b>161</b> |
| 6.1      | Future directions . . . . .    | 164        |
|          | <b>References</b>              | <b>165</b> |

# List of Figures

|    |  |     |
|----|--|-----|
| 1  | Forward-backward splitting as majorization-minimization        | 25  |
| 2  | Forward-backward envelope construction . . . . .               | 27  |
| 3  | Forward-backward envelope bounds . . . . .                     | 28  |
| 4  | Lasso: performance on one instance . . . . .                   | 62  |
| 5  | Lasso: performance on multiple instances . . . . .             | 63  |
| 6  | Group lasso: performance of the algorithms . . . . .           | 66  |
| 7  | Matrix completion: performance of the algorithms . . . . .     | 68  |
| 8  | Image restoration: performance of the algorithms . . . . .     | 69  |
| 9  | Image restoration: original and recovered images . . . . .     | 70  |
| 10 | Spring-mass system: initial position . . . . .                 | 95  |
| 11 | Spring-mass system: convergence of the algorithms . . . . .    | 97  |
| 12 | Spring-mass system: CPU time for closed-loop simulations       | 97  |
| 13 | Spring-mass system: effect of soft-state constraints . . . . . | 98  |
| 14 | Oscillating masses: schematic representation . . . . .         | 133 |
| 15 | Oscillating masses: average and maximum CPU time . . . . .     | 134 |
| 16 | Douglas-Rachford envelope for box QP . . . . .                 | 146 |
| 17 | Box QP: performance of (fast) DRS . . . . .                    | 155 |
| 18 | Lasso: stepsize selection in DRS . . . . .                     | 157 |
| 19 | Analogy between Moreau envelope and FBE/DRE . . . . .          | 162 |

# List of Tables

|   |   |     |
|---|---|-----|
| 2 | Sparse logistic regression: performance of the algorithms . | 65  |
| 3 | Aircraft control: performance of the algorithms . . . . .   | 132 |
| 4 | Comparison between the proposed algorithms . . . . .        | 163 |

## Acknowledgements

I would first like to express my sincere gratitude to Prof. Panagiotis Patrinos: none of this would be possible without his directions, support, and constant motivation and ideas. Most importantly, Panos introduced me to a very exciting area of research, and provided me with tools that I alone would otherwise hardly be able to master. It was an honor and a privilege to closely work with him.

I am also grateful to Prof. Alberto Bemporad for letting me be part of the DYSCO research unit at IMT Lucca: his vast experience and constructive advices have been fundamental throughout these years.

I greatly acknowledge the ESAT/Stadius division of KU Leuven, directed by Prof. Marc Moonen, and all its members and staff with which I had the pleasure to meet, discuss, collaborate. They welcomed me and supported me throughout the final year of PhD, and I am fully indebted with them.

Special thanks go to Andreas for lending me his math skills in numerous occasions, and his  $\text{\TeX}$  typesettings. Thanks to all my close friends and colleagues from IMT Lucca and KU Leuven, with whom I shared more than I could list in a page. Thank you for coffee, drinks, laughter, ping pong, board games, calcetto, LAN parties, NERF guns, crazy talks and late night office madness: (in alphabetical order) Ajay, Alberto, Alessio, Andrea, Andreas, Benedetta, Bertrand, Chris, Daniele, Davide, Duccio, Fabiana, Francesco, Giada, Gionata, Giovanni, Luca, Marco, Michela, Pantelis, Pietro, Puya, Rita, Samuele. It really was a wonderful time.

# Vita

- Dec. 12, 1985** Born in Florence, Italy
- 2004 – 2008** B.Sc. in Computer Science  
Final mark: 110/110  
University of Florence, Italy
- 2008 – 2011** M.Sc. in Computer Science  
Final mark: 110/110 cum laude  
University of Florence, Italy
- 2013 – 2017** Ph.D. in Computer, Decision and Systems Science  
IMT Lucca, Italy
- 09/15 – 01/16** Visiting student, KU Leuven, Belgium  
ESAT – Department of Electrical Engineering
- Since 03/16** Ph.D. student jointly at KU Leuven, Belgium  
ESAT – Department of Electrical Engineering

# Publications

## Journal papers

1. L. Stella, A. Themelis and P. Patrinos. Forward-backward quasi-Newton methods for nonsmooth optimization problems. *Computational Optimization and Applications*, 67(3):443–487, 2017. [10.1007/s10589-017-9912-y](https://doi.org/10.1007/s10589-017-9912-y).
2. L. Stella, A. Themelis and P. Patrinos. Newton-type alternating minimization algorithm for convex optimization. *Submitted*, 2017.
3. A. Themelis, L. Stella and P. Patrinos. Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone line-search algorithms. *arXiv:1606.06256*, 2016.

## Conference proceedings

1. L. Stella, A. Themelis, P. Sopasakis and P. Patrinos. A simple and efficient algorithm for nonlinear model predictive control. *Submitted*, 2017.
2. P. Latafat, L. Stella and P. Patrinos. New primal-dual proximal algorithms for distributed optimization. *55th IEEE Conference on Decision and Control*, 1959–1964, 2016. [10.1109/CDC.2016.7798551](https://doi.org/10.1109/CDC.2016.7798551).
3. P. Patrinos, L. Stella and A. Bemporad. Douglas-Rachford splitting: complexity estimates and accelerated variants. *53rd IEEE Conference on Decision and Control*, 4234–4239, 2014. [10.1109/CDC.2014.7040049](https://doi.org/10.1109/CDC.2014.7040049).

## Working papers

1. P. Patrinos, L. Stella and A. Bemporad. Forward-backward truncated Newton methods for convex composite optimization. *arXiv:1402.6655*, 2014.

# Abstract

Nonsmooth optimization problems arise in an ever-growing number of applications in science and engineering. Proximal (or splitting) algorithms are a general approach to a variety of nonsmooth problems, but as with all first order methods their convergence properties are severely affected by ill conditioning of the problem. In this thesis, an interpretation to proximal algorithms as unconstrained gradient methods over an associated function is provided. Such functions are called *proximal envelopes*, in analogy with the well-known Moreau envelope. Proximal envelopes provide a link between nonsmooth and smooth optimization, and allow for the application of more efficient and robust smooth optimization algorithms to the solution of nonsmooth, possibly constrained problems. We consider the case of the forward-backward and Douglas-Rachford splitting methods. In the first case, based on generalized differentiability properties on the original problem terms, we devise superlinearly convergent line-search algorithms based on quasi-Newton directions, that use the same oracle as the forward-backward splitting; furthermore, the analysis is extended to the case where the dual problem is concerned. In the second case a global convergence rate for the Douglas-Rachford splitting is obtained, while an optimal stepsize selection strategy and an accelerated variant of the method is proposed.



# Notation

|                          |  |
|--------------------------|--|
| $\mathbb{R}$             | The real line.   |
| $\overline{\mathbb{R}}$  | Extended real line $\mathbb{R} \cup \{+\infty\}$ .     |
| $\mathbb{R}_+$           | Non-negative real line.                                |
| $\Gamma_0(E)$            | Proper, closed, convex functions over $E$ .            |
| $C^1(E)$                 | Continuously differentiable functions over $E$ .       |
| $C^2(E)$                 | Twice continuously differentiable functions over $E$ . |
| $C_L^{1,1}(E)$           | Functions over $E$ with $L$ -Lipschitz gradient.       |
| $A \rightrightarrows B$  | Set-valued mapping between sets $A$ and $B$ .          |
| $\delta_S$               | Indicator function of set $S$ .                        |
| $\Pi_S$                  | Projection mapping onto set $S$ .                      |
| $\partial f$             | Subdifferential mapping of $f$ .                       |
| $\text{prox}_{\gamma f}$ | Proximal mapping of $f$ with stepsize $\gamma$ .       |
| $\text{fix } T$          | Set of fixed-points of $T$ .                           |
| $\text{zer } T$          | Set of zeros of $T$ .                                  |
| $A^\top$                 | Transpose of matrix $A$ .                              |
| $A^\dagger$              | Moore-Penrose pseudo-inverse of $A$ .                  |
| $\text{range}(A)$        | Range (image) of $A$ .                                 |
| $\text{null}(A)$         | Null space of $A$ .                                    |
| $\lambda_{\min}(A)$      | Minimum eigenvalue of $A$ .                            |
| $\lambda_{\max}(A)$      | Maximum eigenvalue of $A$ .                            |
| $\sigma_i(A)$            | $i$ -th largest singular value of $A$ .                |
| $DF(x)$                  | Semiderivative of $F$ at $x$ .                         |

|                              |  |
|------------------------------|--|
| $JF(x)$                      | Jacobian of $F$ at $x$ .   |
| $d^2f(x v)$                  | Second epi-derivative of $f$ at $x$ for $v$ .                    |
| $f^*$                        | Convex conjugate function of function $f$ .                      |
| $\mathcal{L}_\gamma$         | Augmented Lagrangian function with parameter $\gamma$ .          |
| $f^\gamma$                   | Moreau envelope of $f$ with parameter $\gamma$ .                 |
| $\varphi_\gamma^{\text{FB}}$ | Forward-backward envelope of $\varphi$ with parameter $\gamma$ . |
| $\varphi_\gamma^{\text{DR}}$ | Douglas-Rachford envelope of $\varphi$ with parameter $\gamma$ . |
| ADMM                         | Alternating direction method of multipliers.                     |
| ALM                          | Augmented Lagrangian method.                                     |
| AMA                          | Alternating minimization algorithm.                              |
| DRE                          | Douglas-Rachford envelope.                                       |
| DRS                          | Douglas-Rachford splitting.                                      |
| FBE                          | Forward-backward envelope.                                       |
| FBS                          | Forward-backward splitting.                                      |
| MPC                          | Model predictive control.  |
| NMPC                         | Nonlinear model predictive control.                              |
| PMA                          | Proximal minimization algorithm.                                 |

# Chapter 1

## Introduction

Optimization problems are ubiquitous in science and engineering. Classically, optimization was regarded as the minimization of smooth functions: while the gradient method serves as a simple basic approach to this class of problems, much more efficient algorithms are obtained by considering directions other than the steepest descent one, for example by exploiting second-order information on the cost function like in Newton-type methods. However, nowadays an ever-growing number of applications fundamentally relies on nonsmoothness to enforce prior knowledge in the problem solution. For this reason new, simple iteration schemes have been designed that usually go by the name of *proximal algorithms* or *splitting methods*. These are essentially fixed-point iterations for solving a nonsmooth, nonlinear system of equations defining the stationary points of the cost function. As such, their iterations are very simple and ideal for embedded applications and large-scale problems.

In this thesis we introduce the concept of *proximal envelopes*: these extend and generalize the concept of *Moreau envelope*, a very well known object in convex optimization, and allow to easily reformulate nonsmooth (possibly constrained) problems as smooth unconstrained ones. As a consequence, new efficient algorithms for certain types of nonsmooth,

nonconvex optimization problem are developed and analyzed. The proposed methods combine the modern approach of proximal algorithms, which are first order methods, with classical techniques for smooth unconstrained optimization that allow for fast asymptotic convergence. The rest of this chapter serves both as introduction for terminology and background notions, as well as to motivate and introduce the idea of proximal envelopes, and to summarize the contributions of the thesis. [Section 1.A](#) contains a summary of the mathematical tools which are used throughout the thesis, and the relative notation.

## 1.1 Structured optimization

In this thesis we are essentially concerned with problems of the form

$$\underset{x}{\text{minimize}} \quad \varphi(x) = f(x) + g(x), \quad (1.1)$$

where  $f$  is a smooth (possibly nonconvex) function, while  $g$  is a convex (possibly nonsmooth) function.

Regularization methods in machine learning, statistics, signal processing, often result in problems of the form (1.1). Here  $f$  can be a smooth penalty, or fitting term, composed with a matrix containing the problem data, while  $g$  is a (possibly nonsmooth) regularization term used to enforce some prior knowledge on the problem solution. For example, the  $\ell_1$  norm  $g(x) = \|x\|_1 = \sum_i |x_i|$  is typically used to enforce sparsity: the *lasso* model for sparse linear regression and the sparse logistic model for classification fit this framework, see [151], [65, §3.4.2, §4.4.4]. The *elastic-net*  $g(x) = \lambda_1 \|x\|_1 + \frac{\lambda_2}{2} \|x\|_2^2$  combines the  $\ell_1$  and Thikonov regularization [164]. If the variable  $x$  is partitioned as  $x = (x_1, \dots, x_N)$ , then  $g(x) = \lambda \sum_{i=1}^N \|x_i\|_2$  is used to enforce *group* sparsity in the solution [65, §3.8.4]. When the decision variable is a matrix  $X$  then the *nuclear norm*  $g(X) = \|X\|_* = \sum_i \sigma_i(X)$  (where  $\sigma_i$  indicates the  $i$ -th singular value) is used to enforce the solution to have low rank, *i.e.*, to be sparse in a

spectral sense [34]. For example, to find a low-rank approximation of a matrix  $M$  from a (noisy) incomplete set  $S$  of measurements of its entries, one can use

$$f(X) = \sum_{(i,j) \in S} (X_{i,j} - M_{i,j})^2, \quad g(X) = \lambda \|X\|_*, \quad (1.2)$$

and problems of this type have applications in collaborative filtering, machine learning, control, remote sensing, and computer vision, see [33, 32] and references therein. See [3] for algorithms with application to sparse and low-rank regularization, and [141] for an extension of this framework to *tensors*. Matrix decomposition problems are formulated as (1.1) with

$$f(X, Y) = \frac{1}{2} \|X + Y - B\|_F^2, \quad g(X, Y) = \lambda \|X\|_1 + \mu \|Y\|_*, \quad (1.3)$$

see [35], and the goal is to approximate matrix  $B$  as the sum of a sparse component  $X$  and a low-rank component  $Y$ . This has applications in video processing for example, where models such as (1.3) are used to perform background subtraction from sequences of frames [159].

A more general form of optimization problems that we consider is

$$\begin{aligned} & \underset{x, z}{\text{minimize}} \quad f(x) + g(z) \\ & \text{subject to} \quad Ax + Bz = b, \end{aligned} \quad (1.4)$$

where  $A, B$  are linear operators and  $b$  is a vector, while  $f$  and  $g$  are functions defined over appropriate spaces. If  $B = -I$  then (1.4) reduces to

$$\underset{x}{\text{minimize}} \quad f(x) + g(Ax - b), \quad (1.5)$$

and problem (1.5) can be used to perform robust regression using a nonsmooth loss function  $g$  such as the  $\ell_1$  norm, the Euclidean norm, or piecewise-linear penalties [25, §6]. The support vector machine model [65,

§12] takes the form (1.5), where  $g(z) = \sum \max\{0, 1 - z\}$  is the hinge loss. In image and signal processing,  $A$  can be a linear operator that computes finite differences, and  $g$  a (group) sparsity inducing penalty, in which case (1.5) is used to denoise a given signal without affecting rapid variations, such as sharp edges in the case of images [134]. Furthermore, if  $g = \delta_{\mathbb{R}_+^m}$  is the indicator function of the nonnegative orthant (see Section 1.A), (1.5) models the minimization of  $f$  over a polyhedral set – a very large class of problems: quadratic programming (QP) falls into this category, and constrained optimal control problems in the framework of *model predictive control* (MPC) very often take this form [114, 144].

Convex problems of the form (1.4) easily allow for decentralized or distributed optimization, when the problem terms are appropriately separable [24]: this is achieved through duality, exploiting the idea of *dual decomposition*. Suppose that  $f(x) = f_1(x_1) + \dots + f_k(x_k)$ , i.e., that  $f$  is separable with respect to a partitioning  $x = (x_1, \dots, x_k)$  of the variable  $x$ . Then it is easy to verify that the dual of (1.4) is

$$\underset{y}{\text{minimize}} \quad f_1^*(A_1^\top y) + \dots + f_k^*(A_k^\top y) + g^*(B^\top y) - \langle b, y \rangle, \quad (1.6)$$

where  $y$  is the dual variable, while  $f_i^*$  and  $g^*$  are the *conjugate functions* of  $f_i$  and  $g$  respectively (see Section 1.A). In this case, when solving (1.6), all quantities associated with  $f_1^*, \dots, f_k^*$  (such as gradients) can be evaluated independently in separate computing nodes, for example distributed over a network. In particular, the data matrices  $A_1, \dots, A_k$  can be handled separately by each computing node, avoiding superfluous (and often impractical due to size, or impossible due to privacy constraints) data exchange. In [24, §7] a variety of patterns that allow for distributed computations is discussed; see [19] for a detailed account on parallel and distributed computations in iterative algorithms. Note that the dual of (1.4) dual takes precisely the form (1.1), which is therefore very comprehensive.

## 1.2 Proximal algorithms

The simplest algorithm to find a solution to (1.1), in the convex case, is arguably the *subgradient method* [140, §2], [18, §3]. From a starting point  $x^0$ , iterate until convergence

$$x^{k+1} = x^k - \gamma_k v^k, \quad v^k \in \partial\varphi(x^k), \quad \gamma_k > 0.$$

This is recognized as a fixed-point iteration of the subdifferential mapping. Due to its simplicity, the subgradient method can be applied to a very wide range of problems. Furthermore, in combination with dual decomposition, this approach results in very simple distributed optimization algorithms [102, 103]. However the simplicity of the method comes at the cost of a usually very slow convergence, and quite restrictive assumptions to converge in the first place: for example, one must consider a diminishing stepsize satisfying  $\sum_k \gamma_k^2 < \infty$ ,  $\sum_k \gamma_k = \infty$  (square-summable but not summable), in order to guarantee convergence under the assumption of bounded subgradients along the iterates [18, Prop. 3.2.6].

Much stronger convergence properties can be obtained by iterating the *resolvent* of  $\partial\varphi$  instead, the so-called *proximal mapping* of  $\varphi$ . In the optimization jargon, this has the following expression [101]

$$\text{prox}_{\gamma\varphi}(x) = \underset{z}{\operatorname{argmin}} \left\{ \varphi(z) + \frac{1}{2\gamma} \|z - x\|^2 \right\}, \quad \gamma > 0.$$

Therefore, evaluating  $\text{prox}_{\gamma\varphi}(x)$  consists of minimizing a regularized version of  $f$  around  $x$ : when  $\varphi$  is convex, such approximation is strongly convex and as a consequence  $\text{prox}_{\gamma\varphi}$  is well defined and has a unique solution. Note that  $\gamma$  here plays the role of a stepsize: when  $\text{prox}_{\gamma\varphi}$  is applied to a point  $x$ , a small  $\gamma$  will yield points which are closer to  $x$ , while a large  $\gamma$  will produce a point which is closer to the minimum of  $\varphi$ . The proximal mapping is easily computable for several functions [112].

If  $S$  is a set and  $\delta_S$  is the indicator function of  $S$ , then  $\text{prox}_{\gamma\delta_S} = \Pi_S$  is the projection onto  $S$ , which is explicitly computable for many types of sets including affine subspaces, halfspaces, boxes,  $\ell_1$  and  $\ell_2$  norm balls, and convex cones such as the nonnegative orthant, second-order cone, positive semidefinite cone, exponential cone. For many other functions the proximal mapping has an explicit expression, such as the  $\ell_1$  norm, (squared)  $\ell_2$  norm, nuclear norm, elastic-net. For convex quadratic functions, evaluating the proximal mapping is a strongly convex quadratic problem and can be either solved exactly using direct methods, or approximately using specialized, very efficient iterative procedures. See [112] for an extensive survey on how proximal mappings can be computed. Iterating the proximal mapping yields the so-called *proximal minimization algorithm* (PMA) [129],

$$x^{k+1} = \text{prox}_{\gamma_k\varphi}(x^k), \quad \gamma_k > 0, \quad (1.7)$$

which is known to converge to a solution with virtually no restriction on the sequence of stepsizes [18, Prop. 5.1.3].

However, for structured problems like (1.1) applying PMA is usually not a trivial task: even if  $\text{prox}_{\gamma f}$  and  $\text{prox}_{\gamma g}$  are easily computable, evaluating  $\text{prox}_{\gamma(f+g)}$  is much harder in general and likely requires an iterative procedure. For this reason, *splitting* algorithms have been proposed that tackle (1.1) by acting separately on  $f$  and  $g$ : in the convex case, such algorithms are the optimization counterpart of operator splitting methods for solving monotone inclusion problems [92, 57]; see [112] for a survey on proximal algorithms and their applications, [9] for an in-depth theoretical analysis of proximal algorithms and their connection to operator splitting methods.

**Forward-backward splitting.** When  $f$  is smooth (*i.e.*, with Lipschitz continuous gradient) and  $g$  has an efficiently computable proximal mapping, one can solve problem (1.1) by alternating gradient (or *forward*)



steps on  $f$  and proximal (or *backward*) steps on  $g$ :

$$x^{k+1} = \text{prox}_{\gamma g}(x^k - \gamma \nabla f(x^k)), \quad \gamma > 0. \quad (1.8)$$

This is known as *proximal gradient method* and is a particular case of the more general *forward-backward splitting* (FBS) [27, 92]. The reason behind this terminology is apparent from the optimality condition of the problem defining the proximal operator: if  $z = \text{prox}_{\gamma g}(x)$ , then necessarily  $z = x - \gamma v$ , with  $v \in \partial g(z)$ , i.e.,  $z$  is obtained by an *implicit* (backward) subgradient step over  $g$ , as opposed to the explicit (forward) step over  $f$ .

In the convex case, FBS is known to converge under minimal assumptions for  $\gamma \in (0, 2/L_f)$ , where  $L_f$  is the Lipschitz constant of  $\nabla f$ , cf. [9, Cor. 27.9]. Moreover, the objective value converges with global sublinear rate and *fast* variants of the method exists with an improved rate [104, 155, 10, 106]. Convergence of FBS has also been shown for nonconvex problems [6]. When applied to the dual of equality constrained convex problems (1.4), FBS results in what is also known as *alternating minimization algorithm* (AMA) [154, 11]. The authors of [112] give an overview of FBS, its interpretations, properties and applications.

**Douglas-Rachford splitting.** Another method for solving (1.1) is the Douglas-Rachford splitting (DRS) [92], which was introduced in the 1950s in the context of the numerical solution of PDEs [53]. This combines proximal steps with respect to both  $f$  and  $g$ , which are assumed to be efficiently computable, as follows:

$$\begin{aligned} y^k &= \text{prox}_{\gamma f}(x^k), \\ z^k &= \text{prox}_{\gamma g}(2y^k - x^k), \\ x^{k+1} &= x^k + \lambda_k(z^k - y^k), \end{aligned} \quad (1.9)$$

where  $\gamma > 0$  and the stepsizes  $\lambda_k \in [0, 2]$  satisfy  $\sum_{k \in \mathbb{N}} \lambda_k(2 - \lambda_k) = +\infty$ . A typical choice for  $\lambda_k$  is to set it equal to 1 for all  $k$ .

If the minimum in (1.1) is attained and the relative interiors of the effective domains of  $f$  and  $g$  have a point in common, it is well known that  $(z^k - y^k)_{k \in \mathbb{N}}$  converges to 0, and  $(x^k)_{k \in \mathbb{N}}$  converges to  $x$  such that  $\text{prox}_{\gamma f}(x) \in \text{argmin } \varphi$  [57, 58, 9]. Therefore  $(y^k)_{k \in \mathbb{N}}$  and  $(z^k)_{k \in \mathbb{N}}$  converge to a solution of (1.1). This general form of DRS was proposed by [57, 58], where it was shown that DRS is a particular case of the proximal point algorithm. Therefore DRS converges under very general assumptions: for example, unlike the forward-backward splitting (1.8), parameter  $\gamma$  can take any positive value. Very recently, sublinear convergence of DRS with respect to the objective value was shown [44]. When applied to the dual of equality constrained convex problems (1.4), DRS can be shown to be equivalent to the *alternating direction method of multipliers* (ADMM), a very well-known algorithm amenable to large-scale problems and distributed optimization [24, 112].

### 1.3 Proximal envelopes

Proximal algorithms, such as FBS and DRS, are first order methods and as such are usually effective at computing low- to medium-precision solutions only. More importantly, their convergence speed is heavily affected by the conditioning properties of the problem at hand [72, 73].

To improve over basic proximal algorithms we will exploit an idea that originates from the following interpretation of the proximal minimization algorithm. The value function of the problem defining  $\text{prox}_{\gamma \varphi}$  is a very well known object in convex optimization, called *Moreau envelope* [101]:

$$\varphi^\gamma(x) = \min_z \left\{ \varphi(z) + \frac{1}{2\gamma} \|z - x\|^2 \right\}.$$

It is not hard to verify that  $\varphi^\gamma(x)$  lower-approximates  $\varphi$  and shares with it its (local) minimizers. Furthermore,  $\varphi^\gamma$  is continuously differentiable with gradient

$$\nabla \varphi^\gamma(x) = \gamma^{-1}(x - \text{prox}_{\gamma \varphi}(x)).$$

Therefore PMA (1.7) can be reformulated as

$$x^{k+1} = x^k - \gamma \nabla \varphi^\gamma(x^k),$$

*i.e.*, it is the gradient method applied to the Moreau envelope. As such, it inherits all properties and drawbacks of the gradient method. When the problem is ill-conditioned, *i.e.*, its solutions lie in a region with very steep and very flat directions (in other words, the problem is badly scaled), then the convergence of gradient methods is known to be severely affected: in the case of  $C^2$  functions, this reflects on ill-conditioning of the Hessian matrix at the problem solution [16, §1.3]. However since PMA is equivalent to the gradient method on the Moreau envelope, which is a smooth function, more advanced iterative schemes can be borrowed from the classical literature of smooth optimization in order to improve its performance. This simple idea provides a link between nonsmooth and smooth optimization and has led to the discovery of a variety of algorithms for problem (1.1), such as semismooth Newton methods [67], variable-metric [22] and quasi-Newton methods [100, 36, 28], and trust-region methods [135]. When PMA is applied to the dual of an equality constrained problem, then the *augmented Lagrangian method* (ALM), also known as *method of multipliers*, is obtained [127, 128]. This fact, in connection with the above considerations, has been exploited to propose Newton-type versions of the method of multipliers [15]. However, evaluating the proximal mapping of  $\varphi = f + g$  (hence the Moreau envelope  $\varphi^\gamma$  and its gradient) is usually nontrivial as we already mentioned: as a consequence, a gradient-based iterative algorithm operating on the Moreau envelope will necessarily require inner iterations to evaluate gradients.

In this thesis we introduce *proximal envelope* functions: these generalize and extend the concept of Moreau envelope to structured problems such as (1.1). In particular, proximal envelopes

- (i) allow to reformulate (1.1), which is nonsmooth and possibly constrained in general, as an equivalent smooth unconstrained prob-

lem (*i.e.*, having the same solutions);

- (ii) offer an interpretation of proximal (splitting) algorithms as gradient methods over a smooth function, suggesting the tempting idea of applying much more efficient and robust algorithms from the classical literature of smooth unconstrained optimization: Newton, quasi-Newton, limited-memory methods [48, 93, 108], nonlinear conjugate gradient methods [71, 42, 43, 76, 77], the Barzilai-Borwein method [7, 64, 125, 41], are all viable approaches to minimize proximal envelope functions, and therefore solve the original nonsmooth problem.

These features render the analogy with the concept of Moreau envelope apparent: in fact, it is easy to show that the Moreau envelope is a particular case of the proximal envelope functions which are discussed here. However, proximal envelopes (and possibly their gradient) can be evaluated at any given point simply using the machinery of the corresponding splitting algorithm. This is in stark contrast with the above mentioned methods based on the Moreau envelope, which usually require an inner iterative procedure.

In particular, we introduce the idea of proximal envelopes in the context of FBS and DRS described above, exploit it to obtain algorithms with improved convergence properties, and demonstrate the practical efficacy of this approach by applying the proposed algorithms to a variety of problems.

In the first case the *forward-backward envelope* (FBE), first proposed in [113], allows us to develop line-search algorithms with global convergence properties and fast (superlinear) asymptotic rate when Newton-type directions are employed [145, 147, 146]. This is achieved through the analysis of first- and second-order properties of the FBE, and observing that these are ensured by mild, generalized differentiability assumptions on the original problem. Other approaches have been taken towards the inclusion of second-order information in the iterations of FBS, to improve

convergence [138, 12, 86, 137]. However, algorithms resulting from these approaches have the limitation that they require an inner iterative procedure. Differently, the algorithms presented here rely solely on evaluations of the FBE (and possibly its gradient), and because of this they are based on the same type of black-box oracle as FBS. In particular, they do not require any inner iterative procedure. The algorithms based on the FBE which are introduced in this thesis were implemented in ForBES, a software package for MATLAB which is available online<sup>1</sup>: this contains generic implementations of the proposed algorithms, and allows to apply them to a variety of applications by specifying the problem terms in (1.1) or (1.4).

In the second case the *Douglas-Rachford envelope* (DRE) allows to derive a sublinear global convergence rate of order  $O(1/k)$  for the objective value (a result that was unknown in general until very recently [44]), a linear rate in the strongly convex case, and an optimal stepsize selection for DRS under appropriate assumptions. Furthermore, an accelerated version of the method is derived, with global rate  $O(1/k^2)$ , when one of the two summands of  $\varphi$  is quadratic [116].

## 1.4 Contributions and organization

The contributions and structure of the thesis are outlined as follows.

Chapter 2, based on [145]: We introduce the forward-backward envelope (FBE), which will be exploited heavily throughout the thesis, and analyzes its properties. In particular, first- and second-order differentiability of the FBE are investigated. A gradient-based line-search algorithm, based on the FBE, are then introduced. The proposed algorithm has similar global convergence properties as FBS (in particular, it enjoys a global sublinear rate in the convex case). Furthermore, superlinear asymptotic convergence of the algorithm is studied when quasi-Newton directions

---

<sup>1</sup><http://kul-forbes.github.io/ForBES/>

are employed in the line-search.

Chapter 3, based on [147]: We consider another algorithmic scheme, based on forward-backward operations and the FBE, which is conceptually simpler than the one in Chapter 2. In particular, no gradient evaluation of the FBE is performed, hence no second order information on the smooth term of the cost is required. Nevertheless, fast directions can be computed that allow to show superlinear convergence of the iterates. The proposed algorithm is applied to nonlinear MPC problems, which are nonconvex.

Chapter 4, based on [146]: Here we are concerned with convex separable problems with linear equality constraints, and the dual problem is tackled. In this case the FBE is shown to be equivalent to the augmented Lagrangian function of the primal problem, evaluated at certain specific primal points. First- and second-order properties of the FBE are extended to this case, and linked directly to generalized differentiability properties of the primal cost. A similar algorithm to the one in Chapter 3 is considered in this context, which is an extension of the alternating minimization algorithm. The algorithm converges superlinearly when quasi-Newton directions are considered, and we also show sublinear and linear convergence rates under mild assumptions.

Chapter 5, based on [116]: We explore the idea of proximal envelopes in the context of the Douglas-Rachford splitting. The analysis is restricted to convex problems, in the case where one term in the objective is smooth: when this is quadratic in particular, then the Douglas-Rachford splitting is equivalent to a gradient method over a smooth convex function. Consequently, a sublinear convergence rate is shown for the algorithm, and linear convergence is proved in the strongly convex case. Finally, an optimal stepsize selection and an accelerated version of the method are proposed.

Chapter 6 contains some final remarks and conclusions, and outlines future research directions.

## 1.A Tools and notation

Throughout the thesis,  $\langle \cdot, \cdot \rangle$  is an inner product over a Euclidean space  $E$  and  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  is the associated norm: the exact nature of  $E$  will be clear from the context. We denote by  $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$  the extended real line. Functions with values in  $\overline{\mathbb{R}}$  are said to be *extended real-valued*. The set of continuously differentiable functions on  $E$  having  $L$ -Lipschitz continuous gradient (also referred to as  $L$ -smooth) is denoted by  $C_L^{1,1}(E)$ .

### 1.A.1 Convex and variational analysis tools

The following are basic notions in convex analysis, and can be found for example in [17] or [132]. A set  $C$  is said to be *convex* if any line between two points in the set lies entirely in the set. Formally,

$$\alpha x + (1 - \alpha)y \in C, \quad \text{for all } x, y \in C.$$

The *affine hull* of a convex set  $S \subseteq E$ , denoted  $\text{aff } S$ , is the intersection of all affine subspaces containing  $S$ ; equivalently, it is the set of affine combinations (*i.e.*, linear combinations with coefficients summing to 1) of points in  $S$ . For a convex set  $C \subseteq E$ , the *relative interior* of  $C$  is the set of points  $x \in C$  for which a sphere  $S$  centered in  $x$  exists, such that  $S \cap \text{aff } C \subset C$ . Note that while a convex set may have empty interior (think of a proper subspace of  $E$ ), the relative interior of a nonempty convex set is always nonempty [17, Prop. 1.3.2].

A function  $f : E \rightarrow \overline{\mathbb{R}}$  is said to be *convex* if

$$f(\alpha x + (1 - \alpha)y) \leq \alpha f(x) + (1 - \alpha)f(y), \quad \text{for all } x, y \in E, \alpha \in (0, 1).$$

An alternative definition can be given in terms of the *epigraph*

$$\text{epi } f = \{(x, v) \mid x \in E, f(x) \leq v\} \subseteq E \times \mathbb{R}.$$

Then  $f$  is convex if and only if  $\text{epi } f$  is a convex set. Another important set associated with extended-real-valued functions is the *effective domain*, i.e., the subset  $\text{dom } f = \{x \in E \mid f(x) < +\infty\} \subseteq E$ . If  $f$  is convex then  $\text{dom } f$  is necessarily a convex set. Furthermore, we say that  $f$  is *proper* if  $\text{dom } f$  is nonempty and  $f$  is finite on  $\text{dom } f$ ; *closed* if  $\text{epi } f$  is a closed set in  $E \times \mathbb{R}$ . The family of proper, closed, convex functions defined on  $E$  with values in  $\overline{\mathbb{R}}$  is indicated as  $\Gamma_0(E)$ . Function  $f$  is said to be *strongly convex* with modulus  $c > 0$  if  $f - \frac{c}{2} \|\cdot\|^2$  is convex.

Given a function  $h$  on  $E$ , the (limiting) *subdifferential*  $\partial h$  of  $h$  is the set-valued mapping [133, Def. 8.3]

$$\partial h(x) = \left\{ v \in E \mid \exists (x^k)_{k \in \mathbb{N}}, (v^k \in \hat{\partial} h(x^k))_{k \in \mathbb{N}} \text{ s.t. } x^k \rightarrow x, v^k \rightarrow v \right\}$$

where

$$\hat{\partial} h(x) = \{v \in E \mid h(z) \geq h(x) + \langle v, z - x \rangle + o(\|z - x\|), \text{ for all } z \in E\}$$

is the *regular subdifferential* of  $h$  at  $x$ . This includes the ordinary gradient in the case of continuously differentiable functions, while for  $g \in \Gamma_0(E)$  it is equivalent to the usual subdifferential for convex functions, i.e.,

$$\partial g(x) = \{v \in E \mid g(y) \geq g(x) + \langle v, y - x \rangle, \forall y \in E\}.$$

We denote by  $\text{zer } \partial h = \{x \in E \mid 0 \in \partial h(x)\}$  the set of *critical points* of function  $h$ . A necessary condition for a point  $x$  to be a local minimizer of  $h$  is that  $x \in \text{zer } \partial \varphi$  [133, Thm. 10.1]. If  $h$  is convex then the condition is also sufficient, and  $x$  is a global minimizer.

The *proximal mapping* [101] of  $h$ , with stepsize  $\gamma > 0$ , is defined as

$$\text{prox}_{\gamma h}(x) = \underset{u \in E}{\text{argmin}} \left\{ h(u) + \frac{1}{2\gamma} \|u - x\|^2 \right\}. \quad (1.10)$$

This can be regarded as a generalized projection, in the sense that if  $\delta_S$  is



the *indicator function* of a nonempty set  $S \subseteq E$ , i.e.,

$$\delta_S(x) = \begin{cases} 0 & \text{if } x \in S, \\ \infty & \text{otherwise,} \end{cases}$$

then  $\text{prox}_{\gamma\delta_S} = \Pi_S$  is the projection onto  $S$  for any  $\gamma > 0$ . The value function of the optimization problem (1.10) defining the proximal mapping is called the *Moreau envelope* [101] and is denoted by  $h^\gamma$ , i.e.,

$$h^\gamma(x) = \min_{u \in E} \left\{ h(u) + \frac{1}{2\gamma} \|u - x\|^2 \right\}. \quad (1.11)$$

Properties of the Moreau envelope and the proximal mapping are well documented in the literature [9, 133, 38, 37]. For a closed, proper function  $h$ , it holds  $h^\gamma \leq h$ , and  $h^\gamma(x) = h(x)$  for any critical point  $x$ . If  $h$  is also convex, then  $\text{prox}_{\gamma h}$  is single-valued, continuous and nonexpansive (with Lipschitz constant 1) and  $h^\gamma$  is convex, continuously differentiable, with gradient

$$\nabla h^\gamma(x) = \gamma^{-1}(x - \text{prox}_{\gamma h}(x)), \quad (1.12)$$

which is  $\gamma^{-1}$ -Lipschitz continuous [9, Prop. 12.29].

For  $h \in \Gamma_0(E)$  we denote by  $h^*$  its *Fenchel conjugate* [61, 132], defined as  $h^*(y) = \sup_x \{\langle x, y \rangle - h(x)\} \in \Gamma_0(E)$ . Properties of conjugate functions are well described for example in [132, 81, 9, 133]. Among these we recall the *Fenchel-Young inequality* [9, Prop. 13.13]

$$\langle x, y \rangle \leq h(x) + h^*(y) \quad \forall x, y \in E, \quad (1.13)$$

with in particular (*conjugate subgradient theorem*, [132, Thm. 23.5])

$$\langle x, y \rangle = h(x) + h^*(y) \Leftrightarrow y \in \partial h(x) \Leftrightarrow x \in \partial h^*(y). \quad (1.14)$$

Moreau identity [9, Thm. 14.3(ii)] relates the proximal mapping of  $h$  and  $h^*$  as follows:

$$y = \text{prox}_{\gamma h}(y) + \gamma \text{prox}_{\gamma^{-1}h^*}(\gamma^{-1}y) \quad \forall y \in E. \quad (1.15)$$

### 1.A.2 Continuity and (generalized) differentiability

We follow the terminology of [133] when referring to the concepts of *strict continuity* and *strict differentiability*. For  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ , we say that  $F$  is *strictly continuous* at  $\bar{x}$  if [133, Def. 9.1(b)]

$$\limsup_{\substack{(x,y) \rightarrow (\bar{x},\bar{x}) \\ x \neq y}} \frac{\|F(y) - F(x)\|}{\|y - x\|} < \infty.$$

If  $F$  is (Frechét) differentiable, we let  $JF : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times n}$  denote the *Jacobian* of  $F$ . When  $m = 1$  we indicate with  $\nabla F = JF^\top$  the *gradient* of  $F$  and with  $\nabla^2 F = J\nabla F^\top$  its *Hessian*, whenever it makes sense. We say that  $F$  is *strictly differentiable* at  $\bar{x}$  if it satisfies the stronger limit [133, Eq. 9(7)]

$$\lim_{\substack{(x,y) \rightarrow (\bar{x},\bar{x}) \\ x \neq y}} \frac{\|F(y) - F(x) - JF(\bar{x})[y - x]\|}{\|y - x\|} = 0.$$

A mapping  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is *positively homogenous* of degree  $p > 0$  if  $G(\alpha x) = \alpha^p G(x)$  for all  $x \in \mathbb{R}^n$  and  $\alpha \geq 0$ , see [133, Def. 13.4]. If omitted, then it is assumed  $p = 1$ . We will indicate by  $DF(x)$  the *semiderivative* of  $F$ , when this exists, according to the following definition (see [133, Thm. 7.21, Eq. 9(6)]; this is sometimes referred to as *B-derivative* [111, 84]).

**Definition 1.A.1** (Semidifferentiability). *A mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is said to be semidifferentiable at a point  $\bar{x} \in \mathbb{R}^n$  if there exists a positively homogeneous mapping  $DF(\bar{x})[\cdot] : \mathbb{R}^n \rightarrow \mathbb{R}^m$  such that*

$$\lim_{x \rightarrow \bar{x}} \frac{\|F(x) - F(\bar{x}) - DF(\bar{x})[x - \bar{x}]\|}{\|x - \bar{x}\|} = 0.$$

It is strictly semidifferentiable at  $\bar{x}$  if the stronger limit holds

$$\lim_{\substack{(x,y) \rightarrow (\bar{x},\bar{x}) \\ x \neq y}} \frac{\|F(y) - F(x) - DF(\bar{x})[y - x]\|}{\|y - x\|} = 0.$$

$DF(\bar{x})$  is called semiderivative of  $F$  at  $\bar{x}$ . If  $F$  is (strictly) semidifferentiable at every point of a set  $S$ , then it is said to be (strictly) semidifferentiable in  $S$ .

When  $F$  is semidifferentiable then  $DF(x)[d]$  is the directional derivative of  $F$  at  $x$  along the direction  $d$ . Note that when  $DF(x)[d]$  is actually linear in  $d$  (instead of just positively homogeneous), then the ordinary notion of (strict) differentiability is recovered. This is the case, for example, when the semiderivative is continuous:

**Proposition 1.A.2** ([111, Thm. 2]). *Suppose that  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is semidifferentiable in a neighborhood of  $\bar{x} \in \mathbb{R}^n$ . Then, the following are equivalent:*

- (a)  $DF(\cdot)[d]$  is continuous in its first argument at  $\bar{x}$  for all  $d \in \mathbb{R}^n$ ;
- (b)  $F$  is strictly semidifferentiable at  $\bar{x}$ ;
- (c)  $F$  is strictly differentiable at  $\bar{x}$ .

The following definition gives a notion of regularity of mappings that will be used in some convergence results, and its natural extension to set-valued mappings (such as the subdifferential mapping), see [52, §1.C] and discussion thereafter, and [52, §3.H, Ex. 3H.4].

**Definition 1.A.3** (Calmness). *A mapping  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is said to be calm at  $\bar{x}$  if*

$$F(x) \in F(\bar{x}) + O(\|x - \bar{x}\|), \quad \forall x \in \mathbb{R}^n.$$

*A set-valued mapping  $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$  is said to be calm at  $\bar{x} \in \mathbb{R}^n$  for  $\bar{y} \in F(\bar{x})$  if there is a neighborhood  $U$  of  $\bar{y}$  such that*

$$F(x) \cap U \subseteq F(\bar{x}) + O(\|x - \bar{x}\|), \quad \forall x \in \mathbb{R}^n.$$

We simply say that  $F$  is calm at  $\bar{x} \in \mathbb{R}^n$  (with no mention of  $\bar{y}$ ) if it is calm at  $\bar{x} \in \mathbb{R}^m$  for all  $\bar{y} \in F(\bar{x})$ .

The analysis of first- and second-order differentiability of proximal envelopes is based on generalized second-order properties of the nonsmooth cost function. These are due to Rockafellar [133, §13], and concern the convergence of the following second-order difference quotient

$$\Delta_\tau^2 f(x|v)[d] = \frac{f(x + \tau d) - f(x) - \tau \langle v, d \rangle}{\frac{1}{2}\tau^2}$$

as  $\tau \searrow 0$ . Specifically, we will consider cases where a function  $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$  is (strictly) twice epi-differentiable according to the following definition (see [133, Def. 13.6], [120]).

**Definition 1.A.4.** *Function  $f$  is said to be twice epi-differentiable at  $x$  for  $v$ , if the second-order difference quotient  $\Delta_\tau^2 f(x|v)$  epi-converges as  $\tau \searrow 0$  (i.e., its epigraph converges in the sense of Painlevé-Kuratowski, see [133, Def. 7.1]), the limit being the function  $d^2 f(x|v)$  given by*

$$d^2 f(x|v)[d] = \liminf_{\substack{\tau \searrow 0 \\ d' \rightarrow d}} \Delta_\tau^2 f(x|v)[d']. \quad (1.16)$$

In this case (1.16), as a function of  $d$ , is said to be the second-order epi-derivative of  $f$  at  $x$  for  $v$ . If  $\Delta_\tau^2 f(\bar{x}|\bar{v})$  epi-converges as  $\tau \searrow 0$ ,  $\bar{x} \rightarrow x$  and  $\bar{v} \rightarrow v$ , then  $f$  is said to be strictly twice epi-differentiable.

Twice epi-differentiability is a mild requirement, and functions with this property are abundant. Refer to [130, 131, 118, 119, 120] and to [133, §7, §13] for examples and an in-depth account on epi-derivatives, epi-differentiability, and their connections with ordinary differentiability.

### 1.A.3 Convergence rates

We will talk about the linear and superlinear convergence of the proposed algorithms according to the following definition (see also [49, Def.

2.3.1] and discussion thereafter).

**Definition 1.A.5.** We say that  $(x^k)_{k \in \mathbb{N}}$  converges to  $x_*$

(i) *Q-linearly with factor  $\omega \in (0, 1)$  if  $\|x^{k+1} - x_*\| \leq \omega \|x^k - x_*\|$  for all  $k \geq 0$ ;*

(ii) *Q-superlinearly if  $\|x^{k+1} - x_*\| / \|x^k - x_*\| \rightarrow 0$ .*

The convergence is *R-linear* (respectively, *R-superlinear*) if  $\|x^k - x_*\| \leq a_k$  for all  $k \geq 0$  and a sequence  $(a_k)_{k \in \mathbb{N}}$  such that  $a_k \rightarrow 0$  with *Q-linear* (respectively, *Q-superlinear*) rate.

Note that linear convergence is sometimes defined in the asymptotic sense, i.e., as  $\|x^{k+1} - x_*\| / \|x^k - x_*\| \rightarrow \omega \in (0, 1)$ . Here we stick to [Definition 1.A.5\(i\)](#) as it allows to capture linear convergence both in the global and local sense (by restricting to tail subsequences).

# Chapter 2

## Forward-backward quasi-Newton methods

### 2.1 Introduction

In this chapter we focus on nonsmooth optimization problems over  $\mathbb{R}^n$  of the form

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad \varphi(x) = f(x) + g(x), \quad (2.1)$$

where  $f$  is a smooth (possibly nonconvex) function, while  $g$  is a proper, closed, convex (possibly nonsmooth) function with cheaply computable proximal mapping [101]. Problems of this form appear in several application fields such as control, system identification, signal and image processing, machine learning and statistics.

Perhaps the most well known algorithm to solve problem (2.1) is the forward-backward splitting (FBS), also known as proximal gradient method [92, 37], which generalizes the classical gradient method to problems involving an additional nonsmooth term. Convergence of the iterates of FBS to a critical point of problem (2.1) has been shown, in the

general nonconvex case, for functions  $\varphi$  having the Kurdyka-Łojasiewicz property [95, 96, 85, 6]. This assumption was used to prove convergence of many other algorithms [4, 5, 6, 21, 110]. The global convergence rate of FBS is known to be sublinear of order  $O(1/k)$  in the convex case, where  $k$  is the iteration count, and can be improved to  $O(1/k^2)$  with techniques based on the work of Nesterov [104, 155, 10, 106]. Therefore, FBS is usually efficient for computing solutions with small to medium precision only and, just like all first order methods, suffers from ill-conditioning of the problem at hand. A remedy to this is to add second-order information in the computation of the forward and backward steps, so to better scale the problem and achieve superlinear asymptotic convergence. As proposed by several authors [12, 86, 137], this can be done by computing the gradient steps and proximal steps according to the  $Q$ -norm rather than the Euclidean norm, where  $Q$  is the Hessian of  $f$  or some approximation to it. This approach has the severe limitation that, unless  $Q$  has a very particular structure, the backward step becomes now very hard and requires an inner iterative procedure to be computed.

Here we follow a different approach. We define a function, which we call *forward-backward envelope* (FBE) that serves as a real-valued, continuously differentiable, exact penalty function for the original problem. Furthermore, forward-backward splitting is shown to be equivalent to a (variable-metric) gradient method applied to the problem of minimizing the FBE. The value and gradient of the FBE can be computed solely based on the evaluation of a forward-backward step at the point of interest. For these reasons, the FBE works as a surrogate of the Moreau envelope [101] for composite problems of the form (2.1). Most importantly, this opens up the possibility of using well-known smooth unconstrained optimization algorithms, with faster asymptotic convergence properties than the gradient method, to minimize the FBE and thus solve (2.1), which is non-smooth and possibly constrained. This approach was first explored in [113], where two Newton-type methods were proposed, and combines and extends ideas stemming from the literature on merit functions for

*variational inequalities* (VIs) and *complementarity problems* (CPs), specifically the reformulation of a VI as a constrained continuously differentiable optimization problem via the regularized gap function [66] and as an unconstrained continuously differentiable optimization problem via the D-gap function [162] (see [59, §10] for a survey and [90, 115] for applications to constrained optimization and model predictive control of dynamical systems).

Then we propose an algorithmic scheme, based on line-search methods, to minimize the FBE. In particular, when descent steps are taken along quasi-Newton directions, superlinear convergence can be achieved when usual nonsingularity assumptions hold at the limit point of the sequence of iterates. The asymptotic analysis is based on an analogous of the Dennis and Moré theorem [47] for the proposed algorithmic scheme, and the BFGS quasi-Newton method is shown to fit this framework. Its limited memory variant L-BFGS, which is suited for large scale problems, is also analyzed. At the same time, we show that our algorithm enjoys the same global convergence properties of FBS under the same assumptions on the original function  $\varphi$ , despite our method operates on the FBE. Unlike the approaches of [12, 86, 137], our algorithm does not require the solution to any inner problem.

The contributions of this chapter can be summarized as follows:

- We give an interpretation of forward-backward splitting as a (variable metric) gradient method over a  $C^1$  function, the forward-backward envelope (FBE). We analyze the fundamental properties of the FBE, including second-order properties around the solutions to (2.1) under mild assumptions on  $g$ .
- We propose an algorithmic scheme for solving problem (2.1) based on line-search methods applied to the problem of minimizing the FBE, and prove that it converges globally to a critical point when  $\varphi$  is convex or has the Kurdyka-Łojasiewicz property. This is a crucial feature of our approach: in fact, the FBE is nonconvex in general,



and there exist examples showing how classical line-search methods need not converge to critical points for nonconvex functions [39, 98, 99, 40]. When  $\varphi$  is convex, in addition, global sublinear convergence of order  $O(1/k)$  (in the objective value) is proved.

- We show that when the directions of choice satisfy the Dennis-Moré condition the method converges superlinearly, under appropriate assumptions, and illustrate when this is the case for BFGS. The resulting algorithm has the same global convergence properties as FBS but, despite relying on the same black-box oracle, converges much faster in practice.

The chapter is organized as follows. [Section 2.2](#) introduces the forward-backward envelope function and illustrates its properties. In [Section 2.3](#) we propose our algorithmic scheme and prove its global convergence properties. Linear convergence is also discussed. [Section 2.4](#) is devoted to the asymptotic convergence analysis in the particular case where quasi-Newton directions are used, specializing the results to the case of BFGS. Limited-memory directions are also discussed. Finally, [Section 2.5](#) illustrates numerical results obtained with the proposed method. Some of the proofs are deferred to the Appendix for the sake of readability.

### 2.1.1 Forward-backward splitting

In the rest of the chapter we will work under the following

**Assumption 2.1.** *In (2.1),  $f \in C_{L_f}^{1,1}(\mathbb{R}^n)$  for  $L_f > 0$  and  $g \in \Gamma_0(\mathbb{R}^n)$ .*

If  $f$  satisfies [Assumption 2.1](#) then [16, Prop. A.24]

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L_f}{2} \|y - x\|^2. \quad (2.2)$$

Given an initial point  $x^0$  and  $\gamma > 0$ , forward-backward splitting (also known as proximal gradient method) seeks solutions to problem (2.1) by

means of the following iterations:

$$x^{k+1} = \text{prox}_{\gamma g}(x^k - \gamma \nabla f(x^k)). \quad (2.3)$$

Under [Assumption 2.1](#) the generated sequence  $(x^k)_{k \in \mathbb{N}}$  satisfies [\[106, eq. \(2.13\)\]](#)

$$\varphi(x^{k+1}) - \varphi(x^k) \leq -\frac{2-\gamma L_f}{2\gamma} \|x^{k+1} - x^k\|^2.$$

If  $\gamma \in (0, 2/L_f)$  and  $\varphi$  is lower bounded, it can be easily inferred that any cluster point  $x$  is stationary for  $\varphi$ , in the sense that it satisfies the necessary condition for optimality  $x \in \text{zer } \partial\varphi$ . The existence of cluster points is ensured if  $(x^k)_{k \in \mathbb{N}}$  remains bounded; due to the monotonic behavior of  $(\varphi(x^k))_{k \in \mathbb{N}}$  for  $\gamma$  in the given range, this condition in turn is guaranteed if  $\varphi$  and the initial point  $x^0$  satisfy the following requirement, which is a standard assumption for nonconvex problems (see e.g. [\[106\]](#)).

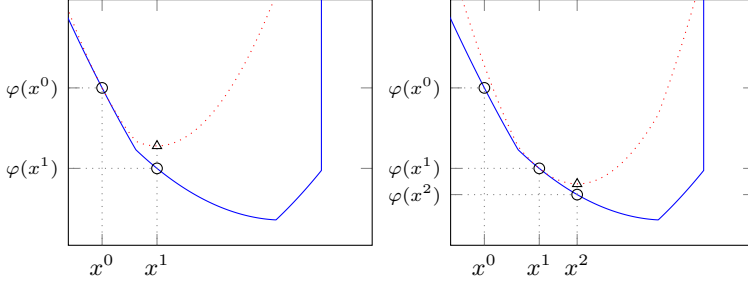
**Assumption 2.2.** *The level set  $\{x \in \mathbb{R}^n \mid \varphi(x) \leq \varphi(x^0)\}$ , which for conciseness we shall denote  $\{\varphi \leq \varphi(x^0)\}$ , is bounded. In particular, there exists  $R > 0$  such that  $\|x - z\| \leq R$  for all  $x \in \{\varphi \leq \varphi(x^0)\}$  and  $z \in \text{argmin } \varphi$ .*

The existence of such a uniform radius  $R$  is due to boundedness of  $\text{argmin } \varphi$ , which in turn follows from the assumed boundedness of the initial level set  $\{\varphi \leq \varphi(x^0)\}$ .

**Example 2.1.1.** To see that  $\text{argmin } \varphi \neq \emptyset$  is not enough for preventing the generation of unbounded sequences, consider  $\varphi = f + g : \mathbb{R} \rightarrow \overline{\mathbb{R}}$  where

$$g = \delta_{(-\infty, 2]} \quad \text{and} \quad f(x) = \begin{cases} e^x - 1 & \text{if } x < 0, \\ x - x^2 & \text{if } x \geq 0. \end{cases}$$

[Assumption 2.1](#) is satisfied with  $L_f = 2$  and  $\text{argmin } \varphi = \{2\}$ . However, for any  $\gamma \in (0, 1)$  the sequence  $(x^k)_{k \in \mathbb{N}}$  generated by [\(2.3\)](#) with  $x^0 < 1/2$  diverges to  $-\infty$ , and  $\varphi(x^k) \rightarrow -1 > -2 = \min \varphi$ . This however cannot happen in the convex case [\[9, Thm. 25.8\]](#).  $\square$



**Figure 1:** When  $\gamma$  is small enough forward-backward splitting minimizes, at every step, a convex majorization (red, dotted lines) of the original cost  $\varphi$  (blue, solid line), cf. (2.7a).

We use shorthands to denote the forward-backward mapping and the associated *fixed-point residual* in order to simplify the notation:

$$T_\gamma(x) = \text{prox}_{\gamma g}(x - \gamma \nabla f(x)), \quad (2.4)$$

$$R_\gamma(x) = \gamma^{-1}(x - T_\gamma(x)), \quad (2.5)$$

so that iteration (2.3) can be written as  $x^{k+1} = T_\gamma(x^k) = x^k - \gamma R_\gamma(x^k)$ . The set  $\text{zer } \partial\varphi$  is easily characterized in terms of the fixed-point set of  $T_\gamma$  as follows:

$$x = T_\gamma(x) \iff x \in \text{zer } \partial\varphi. \quad (2.6)$$

Note that  $T_\gamma(x)$  can alternatively be expressed as the solution to the following partially linearized subproblem (see also Figure 1):

$$T_\gamma(x) = \underset{u \in \mathbb{R}^n}{\text{argmin}} \left\{ \ell_\varphi(u, x) + \frac{1}{2\gamma} \|u - x\|^2 \right\}, \quad (2.7a)$$

$$\ell_\varphi(u, x) = f(x) + \langle \nabla f(x), u - x \rangle + g(u). \quad (2.7b)$$

## 2.2 Forward-backward envelope

We now proceed to the reformulation of (2.1) as the minimization of an unconstrained continuously differentiable function. To this end, we consider the value function of problem (2.7a) defining the forward-backward mapping  $T_\gamma$  and give the following definition.

**Definition 2.2.1** (Forward-backward envelope). *Let  $f, g$  and  $\varphi$  be as in Assumption 2.1, and let  $\gamma > 0$ . The forward-backward envelope (FBE) of  $\varphi$  with parameter  $\gamma$  is*

$$\varphi_\gamma^{\text{FB}}(x) = \min_{u \in \mathbb{R}^n} \left\{ \ell_\varphi(u, x) + \frac{1}{2\gamma} \|u - x\|^2 \right\}. \quad (2.8)$$

Using (2.7a) and (2.7b) it is easy to verify that (2.8) can be equivalently expressed as

$$\varphi_\gamma^{\text{FB}}(x) = f(x) + g(T_\gamma(x)) - \gamma \langle \nabla f(x), R_\gamma(x) \rangle + \frac{\gamma}{2} \|R_\gamma(x)\|^2 \quad (2.9)$$

or, by the definition of Moreau envelope, as

$$\varphi_\gamma^{\text{FB}}(x) = f(x) - \frac{\gamma}{2} \|\nabla f(x)\|^2 + g^\gamma(x - \gamma \nabla f(x)). \quad (2.10)$$

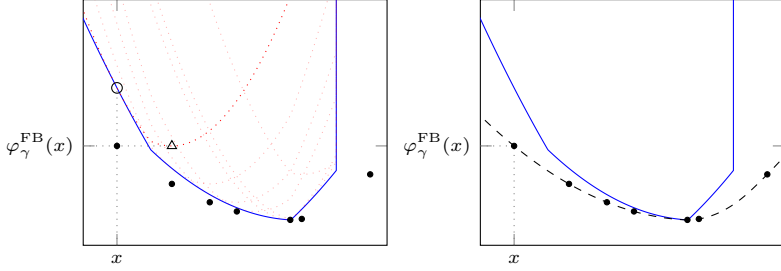
The geometrical construction of  $\varphi_\gamma^{\text{FB}}$  is depicted in Figure 2. One distinctive feature of  $\varphi_\gamma^{\text{FB}}$  is the fact that it is real-valued, despite the fact that  $\varphi$  can be extended-real-valued. Function  $\varphi_\gamma^{\text{FB}}$  has other favorable properties which we now summarize.

### 2.2.1 Basic inequalities

The following result relates the value  $\varphi_\gamma^{\text{FB}}$  with that of  $\varphi$ .

**Proposition 2.2.2.** *Suppose Assumption 2.1 is satisfied. Then  $\varphi_\gamma^{\text{FB}}$  is a strictly continuous function for any  $\gamma > 0$ . Moreover, for all  $x \in \mathbb{R}^n$*

$$(i) \quad \varphi_\gamma^{\text{FB}}(x) \leq \varphi(x) - \frac{\gamma}{2} \|R_\gamma(x)\|^2 \text{ for all } \gamma > 0;$$



**Figure 2:** The forward-backward envelope  $\varphi_\gamma^{\text{FB}}$  (black, dashed line) is obtained by considering the optimal values of problems (2.7a) (dotted lines), and serves as a real-valued lower bound for the original objective  $\varphi$  (blue, solid line).

$$(ii) \quad \varphi(T_\gamma(x)) \leq \varphi_\gamma^{\text{FB}}(x) - \frac{\gamma}{2} (1 - \gamma L_f) \|R_\gamma(x)\|^2 \text{ for all } \gamma > 0;$$

$$(iii) \quad \varphi(T_\gamma(x)) \leq \varphi_\gamma^{\text{FB}}(x) \text{ for all } \gamma \in (0, 1/L_f].$$

*Proof.* From the definition (2.8) and [133, Ex. 10.32], it is apparent that  $\varphi_\gamma^{\text{FB}}$  is strictly continuous.

Regarding 2.2.2(i), from the optimality condition for (2.7a) we have

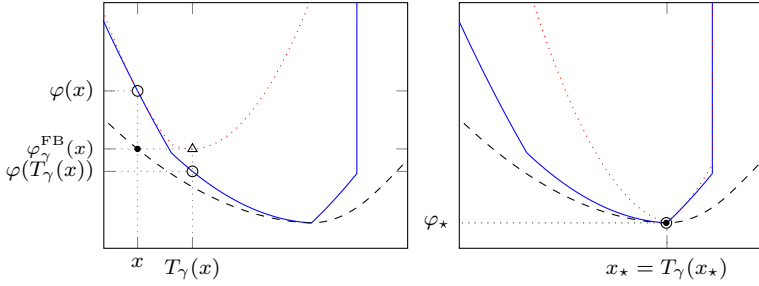
$$R_\gamma(x) - \nabla f(x) \in \partial g(T_\gamma(x)),$$

i.e.,  $R_\gamma(x) - \nabla f(x)$  is a subgradient of  $g$  at  $T_\gamma(x)$ . From subgradient inequality

$$\begin{aligned} g(x) &\geq g(T_\gamma(x)) + \langle R_\gamma(x) - \nabla f(x), x - T_\gamma(x) \rangle \\ &= g(T_\gamma(x)) - \gamma \langle \nabla f(x), R_\gamma(x) \rangle + \gamma \|R_\gamma(x)\|^2. \end{aligned}$$

Adding  $f(x)$  to both sides and considering (2.9) proves the claim. For 2.2.2(ii), we have

$$\begin{aligned} \varphi_\gamma^{\text{FB}}(x) &= f(x) + \gamma \langle \nabla f(x), R_\gamma(x) \rangle + g(T_\gamma(x)) + \frac{\gamma}{2} \|R_\gamma(x)\|^2 \\ &\geq f(T_\gamma(x)) + g(T_\gamma(x)) - \frac{L_f}{2} \|T_\gamma(x) - x\|^2 + \frac{\gamma}{2} \|R_\gamma(x)\|^2 \end{aligned}$$



**Figure 3:** On the left, by [Proposition 2.2.2](#)  $\varphi_\gamma^{\text{FB}}(x)$  is upper bounded by  $\varphi(x)$  and, when  $\gamma$  is small enough, lower bounded by  $\varphi(T_\gamma(x))$ . On the right, by [Proposition 2.2.3\(i\)](#) the two bounds coincide in correspondence of critical points.

where the inequality follows by (2.2). [2.2.2\(iii\)](#) then trivially follows.  $\square$

A consequence of [Proposition 2.2.2](#) is that, whenever  $\gamma$  is small enough, the problems of minimizing  $\varphi$  and  $\varphi_\gamma^{\text{FB}}$  are equivalent.

**Proposition 2.2.3.** *Suppose [Assumption 2.1](#) is satisfied. Then,*

- (i)  $\varphi(z) = \varphi_\gamma^{\text{FB}}(z)$  for all  $\gamma > 0$  and  $z \in \text{zer } \partial\varphi$ ;
- (ii)  $\inf \varphi = \inf \varphi_\gamma^{\text{FB}}$  and  $\text{argmin } \varphi \subseteq \text{argmin } \varphi_\gamma^{\text{FB}}$  for  $\gamma \in (0, 1/L_f]$ ;
- (iii)  $\text{argmin } \varphi = \text{argmin } \varphi_\gamma^{\text{FB}}$  for all  $\gamma \in (0, 1/L_f]$ .

*Proof.* [2.2.3\(i\)](#) follows from (2.6), [Propositions 2.2.2\(i\)](#) and [2.2.2\(ii\)](#).

Suppose now  $\gamma \in (0, 1/L_f]$ . In particular, [2.2.3\(i\)](#) holds for any  $x_* \in \text{argmin } \varphi$ , so

$$\varphi_\gamma^{\text{FB}}(x_*) = \varphi(x_*) \leq \varphi(T_\gamma(x)) \leq \varphi_\gamma^{\text{FB}}(x) \quad \text{for all } x \in \mathbb{R}^n$$

where the first inequality follows from optimality of  $x_*$  for  $\varphi$ , and the second from [Proposition 2.2.2\(iii\)](#). Therefore, every  $x_* \in \text{argmin } \varphi$  is also a minimizer of  $\varphi_\gamma^{\text{FB}}$ , and  $\min \varphi = \min \varphi_\gamma^{\text{FB}}$  provided that the former is attained. It remains to show the case  $\text{argmin } \varphi = \emptyset$ . By [Proposition 2.2.2\(i\)](#) we have  $\inf \varphi_\gamma^{\text{FB}} \leq \inf \varphi$ . If there exists  $x \in \mathbb{R}^n$  such that

$\varphi_\gamma^{\text{FB}}(x) \leq \inf \varphi$ , then [Proposition 2.2.2\(ii\)](#) implies that  $\varphi(T_\gamma(x)) \leq \inf \varphi$ , contradicting  $\text{argmin } \varphi = \emptyset$ . Therefore  $\inf \varphi_\gamma^{\text{FB}} = \inf \varphi$ , proving [2.2.3\(ii\)](#).

Suppose now  $\gamma \in (0, 1/L_f)$ , and let  $x_\star \in \text{argmin } \varphi_\gamma^{\text{FB}}$ . From [Propositions 2.2.2\(i\)](#) and [2.2.2\(ii\)](#) we get that

$$\varphi_\gamma^{\text{FB}}(T_\gamma(x_\star)) \leq \varphi(T_\gamma(x_\star)) \leq \varphi_\gamma^{\text{FB}}(x_\star) - \frac{1-\gamma L_f}{2} \|x_\star - T_\gamma(x_\star)\|^2,$$

which implies  $x_\star = T_\gamma(x_\star)$ , since  $x_\star$  minimizes  $\varphi_\gamma^{\text{FB}}$  and  $\frac{1-\gamma L_f}{2} > 0$ . Therefore, the following chain of inequalities holds

$$\varphi_\gamma^{\text{FB}}(x_\star) = \varphi_\gamma^{\text{FB}}(T_\gamma(x_\star)) \leq \varphi(x_\star) \leq \varphi_\gamma^{\text{FB}}(x_\star).$$

Since  $\varphi_\gamma^{\text{FB}} \leq \varphi$  and  $x_\star$  minimizes  $\varphi_\gamma^{\text{FB}}$ , it follows that  $x_\star \in \text{argmin } \varphi$ . Therefore, the sets of minimizers of  $\varphi$  and  $\varphi_\gamma^{\text{FB}}$  coincide, proving [2.2.3\(iii\)](#).  $\square$

**Example 2.2.4.** To see that the bounds on  $\gamma$  in [Proposition 2.2.3](#) are tight, consider the convex problem

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \varphi(x) = \sqrt{\frac{f(x)}{2} \|x\|^2} + \sqrt{\delta_{\mathbb{R}_+^n}(x)}$$

where  $\mathbb{R}_+^n = \{x \in \mathbb{R}^n \mid x_i \geq 0, i = 1 \dots n\}$  is the nonnegative orthant. [Assumption 2.1](#) is satisfied with  $L_f = 1$ , and the only stationary point for  $\varphi$  is the unique minimizer  $x_\star = 0$ . Using [\(2.10\)](#) we can explicitly compute the FBE: for any  $\gamma > 0$  we have

$$\varphi_\gamma^{\text{FB}}(x) = \frac{1-\gamma}{2} \|x\|^2 + \frac{1}{2\gamma} \|(1-\gamma)x - [(1-\gamma)x]_+\|^2,$$

where  $[x]_+ = \Pi_{\mathbb{R}_+^n}(x) = \max\{x, 0\}$ , the last expression being meant componentwise. For any  $\gamma > 0$  we have that  $\varphi_\gamma^{\text{FB}}(x_\star) = \varphi(x_\star)$ , as ensured by [Proposition 2.2.3\(i\)](#), and as long as  $\gamma < 1 = 1/L_f$  all properties in [Proposition 2.2.3](#) do hold. For  $\gamma = 1$  we have that  $\varphi_\gamma^{\text{FB}} \equiv 0$ , showing the inclusion in [Proposition 2.2.3\(ii\)](#) to be proper, yet satisfying  $\min \varphi_\gamma^{\text{FB}} = \min \varphi$ .

However, for  $\gamma > 1$  the FBE  $\varphi_\gamma^{\text{FB}}$  is not even lower bounded, as it can be easily deduced by observing that, letting  $x^k = (-k, 0, \dots, 0)$  for  $k \in \mathbb{N}$ ,  $\varphi_\gamma^{\text{FB}}(x^k) = \frac{1-\gamma}{2}k^2$  is arbitrarily negative.  $\square$

[Proposition 2.2.3](#) implies, using [Proposition 2.2.2\(i\)](#), that an  $\varepsilon$ -optimal solution  $x$  of  $\varphi$  is automatically  $\varepsilon$ -optimal for  $\varphi_\gamma^{\text{FB}}$  and, using [Proposition 2.2.2\(ii\)](#), from an  $\varepsilon$ -optimal solution for  $\varphi_\gamma^{\text{FB}}$  we can directly obtain an  $\varepsilon$ -optimal solution for  $\varphi$  if  $\gamma \in (0, 1/L_f]$ :

$$\begin{aligned}\varphi(x) - \inf \varphi \leq \varepsilon &\implies \varphi_\gamma^{\text{FB}}(x) - \inf \varphi \leq \varepsilon \\ \varphi_\gamma^{\text{FB}}(x) - \inf \varphi_\gamma^{\text{FB}} \leq \varepsilon &\implies \varphi(T_\gamma(x)) - \inf \varphi \leq \varepsilon.\end{aligned}$$

[Proposition 2.2.3](#) also highlights the first apparent similarity between the concepts of FBE and Moreau envelope (1.11): the latter is indeed itself a lower bound for the original function, sharing with it its minimizers and minimum value. In fact, the two are directly related as we now show. In particular, the following result implies that if  $\varphi$  is convex (e.g. if  $f$  is) and  $\gamma \in (0, 1/L_f)$ , then the possibly nonconvex  $\varphi_\gamma^{\text{FB}}$  is upper and lower bounded by convex functions.

**Proposition 2.2.5.** *Suppose [Assumption 2.1](#) is satisfied. Then,*

- (i)  $\varphi_\gamma^{\text{FB}} \leq \varphi^{\frac{\gamma}{1+\gamma L_f}}$  for all  $\gamma > 0$ ;
- (ii)  $\varphi^{\frac{\gamma}{1-\gamma L_f}} \leq \varphi_\gamma^{\text{FB}}$  for all  $\gamma \in (0, 1/L_f)$ ;
- (iii)  $\varphi_\gamma^{\text{FB}} \leq \varphi^\gamma$  if  $f$  is convex.

*Proof.* (2.2) implies the following bounds concerning the partial linearization:

$$-\frac{L_f}{2}\|u - x\|^2 \leq \varphi(u) - \ell_\varphi(u, x) \leq \frac{L_f}{2}\|u - x\|^2.$$

Combined with the definition of the FBE, cf. (2.8), this proves [2.2.5\(i\)](#) and [2.2.5\(ii\)](#).

If  $f$  is convex, the lower bound can be strengthened to  $0 \leq \varphi(u) - \ell_\varphi(u, x)$ . Adding  $\frac{1}{2\gamma}\|u - x\|^2$  to both sides and minimizing with respect



to  $u$  yields 2.2.5(iii). □

## 2.2.2 First- and second-order properties

We now turn our attention to differentiability of  $\varphi_\gamma^{\text{FB}}$ , which is fundamental in devising and analyzing algorithms for solving (2.1). To ensure continuous differentiability of  $\varphi_\gamma^{\text{FB}}$  we will need the following

**Assumption 2.3.** *Function  $f$  is twice-continuously differentiable.*

Under Assumption 2.3, the function

$$Q_\gamma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n} \quad \text{given by} \quad Q_\gamma(x) = I - \gamma \nabla^2 f(x) \quad (2.11)$$

is well defined, continuous, and symmetric-valued.

**Theorem 2.2.6** (Differentiability of  $\varphi_\gamma^{\text{FB}}$ ). *Suppose that Assumptions 2.1 and 2.3 are satisfied. Then,  $\varphi_\gamma^{\text{FB}}$  is continuously differentiable with*

$$\nabla \varphi_\gamma^{\text{FB}}(x) = Q_\gamma(x) R_\gamma(x). \quad (2.12)$$

If  $\gamma \in (0, 1/L_f)$  then the set of stationary points of  $\varphi_\gamma^{\text{FB}}$  equals  $\text{zer } \partial \varphi$ .

*Proof.* Consider expression (2.10) for  $\varphi_\gamma^{\text{FB}}$ . The gradient of  $g^\gamma$  is given by (1.12), and since  $f \in C^2$  we have

$$\begin{aligned} \nabla \varphi_\gamma^{\text{FB}}(x) &= \nabla f(x) - \gamma \nabla^2 f(x) \nabla f(x) \\ &\quad + \gamma^{-1} (I - \gamma \nabla^2 f(x)) (x - \gamma \nabla f(x) - T_\gamma(x)) \\ &= (I - \gamma \nabla^2 f(x)) (\nabla f(x) - \nabla f(x) + \gamma^{-1}(x - T_\gamma(x))). \end{aligned}$$

This proves (2.12). If  $\gamma \in (0, 1/L_f)$  then  $Q_\gamma(x)$  is nonsingular for all  $x$ , and therefore  $\nabla \varphi_\gamma^{\text{FB}}(x) = 0$  if and only if  $R_\gamma(x) = 0$ , which means that  $x$  is a critical point of  $\varphi$  by (2.6). □

Together with Proposition 2.2.3, Theorem 2.2.6 shows that if  $\gamma \in (0, 1/L_f)$  the nonsmooth problem (2.1) is completely equivalent to the unconstrained

minimization of the continuously differentiable function  $\varphi_\gamma^{\text{FB}}$ , in the sense that the sets of minimizers and optimal values are equal. In particular, as remarked in the next statement, if  $\varphi$  is convex then the set of stationary points of  $\varphi_\gamma^{\text{FB}}$  turns out to be equal to the set of its minimizers, even though  $\varphi_\gamma^{\text{FB}}$  may not be convex.

**Corollary 2.2.7.** *Suppose that [Assumptions 2.1](#) and [2.3](#) are satisfied. If  $\varphi$  is convex (e.g. if  $f$  is), then  $\text{argmin } \varphi = \text{zer } \nabla \varphi_\gamma^{\text{FB}}$  for all  $\gamma \in (0, 1/L_f)$ .*

The FBE is not everywhere twice continuously differentiable in general. For example, if  $g$  is real valued then  $g^\gamma \in C^2$  if and only if  $g \in C^2$  [87]. However, second order properties will only be needed at critical points of  $\varphi$  in our framework, and for this purpose we can rely on generalized second-order differentiability notions described in [133, §13].

**Assumption 2.4.** *Function  $g$  is (strictly) twice epi-differentiable at  $x \in \text{zer } \partial\varphi$  for  $-\nabla f(x)$ , with generalized-quadratic second order epi-derivative. That is,*

$$d^2g(x|-\nabla f(x))[d] = \langle d, Md \rangle + \delta_S(d), \quad \forall d \in \mathbb{R}^n \quad (2.13)$$

where  $S \subseteq \mathbb{R}^n$  is a linear subspace, and  $M \in \mathbb{R}^{n \times n}$  is symmetric, positive semidefinite, and such that  $\text{Im}(M) \subseteq S$  and  $\text{Ker}(M) \supseteq S^\perp$ . We refer to the case where  $g$  is strictly twice epi-differentiable as [Assumption 2.4<sup>+</sup>](#).

In some results we will need to assume the following slightly stronger property. The properties of  $M$  in [Assumption 2.4](#) cause no loss of generality. Indeed, letting  $\Pi_S$  denote the orthogonal projection onto  $S$  ( $\Pi_S$  is symmetric, see [14]), if matrix  $M \succeq 0$  satisfies (2.13) so does matrix  $M' = \Pi_S[\frac{1}{2}(M + M^\top)]\Pi_S$ , which has the required properties.

Twice epi-differentiability of  $g$  is a mild requirement, and cases where  $d^2g$  is actually generalized quadratic are abundant [130, 131, 118, 119]. For example, if  $g$  is piecewise linear and  $x \in \text{zer } \partial\varphi$ , then from [130, Thm. 3.1] it follows that (2.13) holds if and only if the normal cone

$N_{\partial g(x)}(-\nabla f(x))$  is a linear subspace, which is equivalent to

$$-\nabla f(x) \in \text{relint } \partial g(x)$$

where  $\text{relint } \partial g(x)$  is the relative interior of the convex set  $\partial g(x)$ .

**Example 2.2.8 (Lasso).** Let  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^m$  and  $\lambda > 0$ . Consider  $f(x) = \frac{1}{2} \|Ax - b\|^2$  and  $g(x) = \lambda \|x\|_1$ . Minimizing  $\varphi = f + g$  is a frequent problem known as lasso, and attempts to find a sparse least squares solution to the linear system  $Ax = b$ . One has

$$[\partial g(x)]_i = \begin{cases} \{\lambda\} & x_i > 0 \\ \{-\lambda\} & x_i < 0 \\ [-\lambda, \lambda] & x_i = 0. \end{cases}$$

In this case  $d^2 g(x | -\nabla f(x))$  is generalized quadratic at a solution  $x$  as long as whenever  $x_i = 0$  it holds that  $|(A^T(Ax - b))_i| \neq \lambda$ .  $\square$

We begin by investigating differentiability of the residual mapping  $R_\gamma$ .

**Lemma 2.2.9.** Suppose that [Assumptions 2.1](#) and [2.3](#) are satisfied, and that  $g$  satisfies [Assumption 2.4 \(2.4<sup>+</sup>\)](#) at a point  $x \in \text{zer } \partial \varphi$ . Then,  $\text{prox}_{\gamma g}$  is (strictly) differentiable at  $x - \gamma \nabla f(x)$ , and  $R_\gamma$  is (strictly) differentiable at  $x$  with Jacobian

$$JR_\gamma(x) = \gamma^{-1}(I - P_\gamma(x)Q_\gamma(x)), \quad (2.14)$$

where  $Q_\gamma$  is as in [\(2.11\)](#), and

$$P_\gamma(x) = J \text{prox}_{\gamma g}(x - \gamma \nabla f(x)) = \Pi_S[I + \gamma M]^{-1} \Pi_S. \quad (2.15)$$

Moreover,  $Q_\gamma(x)$  and  $P_\gamma(x)$  are symmetric,  $P_\gamma(x) \succeq 0$ ,  $\|P_\gamma(x)\| \leq 1$ , and if  $\gamma \in (0, 1/L_f)$  then  $Q_\gamma(x) \succ 0$ .

*Proof.* We know from [\[120, Thms. 3.8, 4.1\]](#) that  $\text{prox}_{\gamma g}$  is (strictly) differentiable at  $x - \gamma \nabla f(x)$  if and only if  $g$  satisfies [Assumption 2.4 \(2.4<sup>+</sup>\)](#)

at  $x$  for  $-\nabla f(x)$ . Since  $f \in C^2$  by assumption, then in particular  $\nabla f$  is strictly differentiable. The formula (2.14) follows from [Proposition 2.A.1](#) with  $P = \text{prox}_{\gamma g}$  and  $F(x) = x - \gamma \nabla f(x)$ .

Matrix  $Q_\gamma(x)$  is symmetric since  $f \in C^2$  and positive definite if  $\gamma < 1/L_f$ . To obtain an expression for  $P_\gamma(x) = J \text{prox}_{\gamma g}(x - \gamma \nabla f(x))$  we can apply [\[133, Ex. 13.45\]](#) to the *tilted* function  $g + \langle \nabla f(x), \cdot \rangle$  so that, letting  $d^2 g = d^2 g(x | -\nabla f(x))[\cdot]$  and  $\Pi_S$  the idempotent and symmetric projection matrix on  $S$ ,

$$\begin{aligned} P_\gamma(x)d &= \text{prox}_{(\gamma/2)d^2 g}(d) \\ &= \underset{d' \in S}{\text{argmin}} \left\{ \frac{1}{2} \langle d', M d' \rangle + \frac{1}{2\gamma} \|d' - d\|^2 \right\} \\ &= \Pi_S \underset{d' \in \mathbb{R}^n}{\text{argmin}} \left\{ \frac{1}{2} \langle \Pi_S d', M \Pi_S d' \rangle + \frac{1}{2\gamma} \|\Pi_S d' - d\|^2 \right\} \\ &= \Pi_S (\Pi_S [I + \gamma M] \Pi_S)^\dagger \Pi_S d \\ &= \Pi_S [I + \gamma M]^{-1} \Pi_S d \end{aligned}$$

where  $^\dagger$  indicates the pseudo-inverse, and last equality is due to [\[14, Facts 6.4.12\(i\)-\(ii\) and 6.1.6\(xxxii\)\]](#) and the properties of  $M$  as stated in [Assumption 2.4](#). Clearly  $P_\gamma(x) \succeq 0$  is symmetric and  $\|P_\gamma(x)\| \leq 1$ .  $\square$

Next, we see that differentiability of the residual  $R_\gamma$  is equivalent to that of  $\nabla \varphi_\gamma^{\text{FB}}$ . Mild additional assumptions on  $f$  extend this kinship to strict differentiability. Moreover, all strong (local) minimizers of the original problem, *i.e.*, of  $\varphi$ , are also strong (local) minimizers of  $\varphi_\gamma^{\text{FB}}$  (and vice versa, due to the lower-bound property of  $\varphi_\gamma^{\text{FB}}$ ).

**Theorem 2.2.10.** *Suppose that [Assumptions 2.1](#) and [2.3](#) are satisfied, and that  $g$  satisfies [Assumption 2.4](#) at a point  $x \in \text{zer } \partial \varphi$ . Then,  $\varphi_\gamma^{\text{FB}}$  is twice differentiable at  $x$ , with symmetric Hessian given by*

$$\nabla^2 \varphi_\gamma^{\text{FB}}(x) = \gamma^{-1} Q_\gamma(x) (I - P_\gamma(x) Q_\gamma(x)), \quad (2.16)$$

where  $Q_\gamma$  and  $P_\gamma$  are as in [Lemma 2.2.9](#). If moreover  $\nabla^2 f$  is strictly continuous

at  $x$  and  $g$  satisfies [Assumption 2.4<sup>+</sup>](#) at  $x$ , then  $\varphi_\gamma^{\text{FB}}$  is strictly twice differentiable at  $x$ .

*Proof.* Recall from (2.12) that  $\nabla \varphi_\gamma^{\text{FB}}(x) = Q_\gamma(x)R_\gamma(x)$ . The result follows from [Lemma 2.2.9](#) and [Proposition 2.A.2](#) with  $Q = Q_\gamma$  and  $R = R_\gamma$ .  $\square$

**Theorem 2.2.11.** *Suppose that [Assumptions 2.1](#) and [2.3](#) are satisfied, and that  $g$  satisfies [Assumption 2.4](#) at a point  $x \in \text{zer } \partial\varphi$ . Then, for all  $\gamma \in (0, 1/L_f)$  the following are equivalent:*

- (a)  $x$  is a strong local minimum for  $\varphi$ ;
- (b) for all  $d \in S$ ,  $\langle d, (\nabla^2 f(x) + M)d \rangle > 0$ ;
- (c)  $JR_\gamma(x)$  is similar to a symmetric and positive definite matrix;
- (d)  $\nabla^2 \varphi_\gamma^{\text{FB}}(x) \succ 0$ ;
- (e)  $x$  is a strong local minimum for  $\varphi_\gamma^{\text{FB}}$ .

*Proof.* It follows from [Theorem 2.2.10](#) that the Hessian  $\nabla^2 \varphi_\gamma^{\text{FB}}(x)$  exists and is symmetric. Moreover, from [133, Ex. 13.18] we know that for all  $d \in \mathbb{R}^n$

$$\begin{aligned} d^2 \varphi(x|0)[d] &= \langle d, \nabla^2 f(x)d \rangle + d^2 g(x|-\nabla f(x))[d] \\ &= \langle d, \nabla^2 f(x)d \rangle + \langle d, Md \rangle + \delta_S(d). \end{aligned} \quad (2.17)$$

[2.2.11\(a\)](#)  $\Leftrightarrow$  [2.2.11\(b\)](#): Follows from (2.17), using [133, Thm. 13.24(c)].

[2.2.11\(c\)](#)  $\Leftrightarrow$  [2.2.11\(d\)](#): Letting  $Q = Q_\gamma(x)$ , we see from (2.14) and (2.16) that  $JR_\gamma(x)$  is similar to  $Q^{-1/2} \nabla^2 \varphi_\gamma^{\text{FB}}(x) Q^{-1/2}$ , which is symmetric, and which is positive definite if and only if  $\nabla^2 \varphi_\gamma^{\text{FB}}(x)$  is.

[2.2.11\(b\)](#)  $\Leftrightarrow$  [2.2.11\(c\)](#): From the point above we know that  $JR_\gamma(x)$  has all real eigenvalues, and it can be easily seen to be similar to  $\gamma^{-1}(I - QP)$ , where  $P = P_\gamma(x)$ . We can use [83, Theorem 7.7.3] to see that

$$\lambda_{\min}(I - QP) > 0 \quad \text{if and only if} \quad Q^{-1} \succ P.$$

For all  $d \in S$ , using (2.15) we have then

$$\begin{aligned} \langle d, (Q^{-1} - P)d \rangle &= \langle d, Q^{-1}d \rangle - \langle d, \Pi_S[I + \gamma M]^{-1}\Pi_S d \rangle \\ &= \langle d, Q^{-1}d \rangle - \langle \Pi_S d, [I + \gamma M]^{-1}\Pi_S d \rangle \\ &= \langle d, Q^{-1}d \rangle - \langle d, [I + \gamma M]^{-1}d \rangle \end{aligned}$$

and last quantity is positive if and only if  $I + \gamma M \succ Q$  on  $S$ . By definition of  $Q$ , we then have that this holds if and only if  $\nabla^2 f(x) + M \succ 0$  on  $S$ , which is 2.2.11(b).

2.2.11(d)  $\Leftrightarrow$  2.2.11(e): Trivial since  $\nabla^2 \varphi_\gamma^{\text{FB}}(x)$  exists.  $\square$

### 2.2.3 Interpretations

An interesting observation is that the FBE provides a link between gradient methods and FBS, just like the Moreau envelope (1.11) does for the proximal point algorithm [129]. To see this, consider the problem

$$\text{minimize } g(x) \tag{2.18}$$

where  $g \in \Gamma_0(\mathbb{R}^n)$ . The proximal point algorithm for solving (2.18) is

$$x^{k+1} = \text{prox}_{\gamma g}(x^k). \tag{2.19}$$

It is well known that the proximal point algorithm can be interpreted as a gradient method for minimizing the Moreau envelope of  $g$ , cf. (1.11). Indeed, due to (1.12), iteration (2.19) can be expressed as

$$x^{k+1} = x^k - \gamma \nabla g^\gamma(x^k).$$

This simple idea provides a link between nonsmooth and smooth optimization and has led to the discovery of a variety of algorithms for problem (2.18), such as semismooth Newton methods [67], variable-metric [22] and quasi-Newton methods [100, 36, 28], and trust-region methods [135].

However, when dealing with composite problems, even if  $\text{prox}_{\gamma f}$  and  $\text{prox}_{\gamma g}$  are cheaply computable, computing the proximal mapping of  $\varphi = f + g$  is usually as hard as solving (2.1) itself. On the other hand, forward-backward splitting takes advantage of the structure of the problem by operating separately on the two summands, cf. (2.3). The question that naturally arises is the following:

*Is there a continuously differentiable function that provides an interpretation of FBS as a gradient method, just like the Moreau envelope does for the proximal point algorithm?*

The forward-backward envelope provides an affirmative answer. Specifically, whenever  $f$  is  $C^2$ , FBS can be interpreted as the following (variable-metric) gradient method on the FBE:

$$x^{k+1} = x^k - \gamma(I - \gamma \nabla^2 f(x^k))^{-1} \nabla \varphi_{\gamma}^{\text{FB}}(x^k), \quad (2.20)$$

cf. [Theorem 2.2.6](#). Furthermore, the following properties hold for the Moreau envelope

$$g^{\gamma} \leq g, \quad \inf g^{\gamma} = \inf g, \quad \operatorname{argmin} g^{\gamma} = \operatorname{argmin} g,$$

which correspond to [Propositions 2.2.2\(i\)](#) and [2.2.3](#) for the FBE. The relationship between Moreau envelope and forward-backward envelope is then apparent. This opens the possibility of extending FBS and devising new algorithms for problem (2.1) by simply reconsidering and appropriately adjusting methods for unconstrained minimization of continuously differentiable functions, the most well studied problem in optimization.

## 2.3 Forward-backward line-search methods

We consider line-search methods applied to the problem of minimizing  $\varphi_{\gamma}^{\text{FB}}$ , hence solving (2.1). Requirements of such methods are often restrictive, including convexity or even strong convexity of the objective

---

**Algorithm 1** MinFBE
 

---

**Inputs:**  $x^0 \in \mathbb{R}^n$ ,  $\gamma_0 > 0$ ,  $\sigma \in (0, 1)$ ,  $\beta \in [0, 1)$

**Initialize:**  $k = 0$

- 1: Select  $d^k$  such that  $\langle d^k, \nabla \varphi_{\gamma_k}^{\text{FB}}(x^k) \rangle \leq 0$
- 2: Select  $\tau_k \geq 0$  and set  $w^k = x^k + \tau_k d^k$  such that  $\varphi_{\gamma_k}^{\text{FB}}(w^k) \leq \varphi_{\gamma_k}^{\text{FB}}(x^k)$
- 3: If

$$f(T_{\gamma_k}(w^k)) > f(x^k) - \gamma_k \langle \nabla f(x^k), R_{\gamma_k}(x^k) \rangle + \frac{(1-\beta)\gamma_k}{2} \|R_{\gamma_k}(x^k)\|^2$$

then  $\gamma_k \leftarrow \sigma \gamma_k$  and go to step 1

- 4:  $x^{k+1} = T_{\gamma_k}(w^k)$ ,  $\gamma_{k+1} = \gamma_k$ ,  $k \leftarrow k + 1$ , go to step 1
- 

function, properties that unfortunately the FBE does not satisfy in general. As opposed to this, FBS possesses strong convergence properties and complexity estimates. We now show that it is possible to exploit the composite structure of (2.1) and devise line-search methods with the same global convergence properties and oracle information as FBS.

[Algorithm 1](#), which we call MinFBE, interleaves descent steps over the FBE with forward-backward steps. In particular, steps 1 and 2 provide fast asymptotic convergence when directions  $d^k$  are appropriately selected, while step 4 ensures global convergence: this is of central importance, as such properties are not usually enjoyed by standard line-search methods employed to minimize general nonconvex functions [39, 98, 99, 40]. Moreover, in the convex case we are able to show global convergence rate results which are not typical for line-search methods with e.g. quasi-Newton directions. We anticipate some of the favorable properties that [Algorithm 1](#) shares with FBS:

- square-summability of the residuals for lower bounded  $\varphi$  ([Proposition 2.3.4](#));
- global sublinear rate of the objective for convex  $\varphi$  with bounded level sets ([Theorem 2.3.6](#));
- global convergence when  $\varphi$  has bounded level sets and satisfies the



Kurdyka-Łojasiewicz at its stationary points ([Theorem 2.3.10](#));

- local linear rate when  $\varphi$  has the Łojasiewicz property at its critical points ([Theorem 2.3.11](#)).

Moreover, unlike ordinary line-search methods applied to  $\varphi_\gamma^{\text{FB}}$ , we will see in [Proposition 2.3.4](#) that [Algorithm 1](#) is a descent method both for  $\varphi_\gamma^{\text{FB}}$  and  $\varphi$ . Note that, despite the fact that the algorithm operates on  $\varphi_\gamma^{\text{FB}}$ , all the above properties require assumptions or provide results on  $\varphi$ , i.e., on the original problem.

The parameter  $\gamma$  defining the FBE is adjusted in [step 3](#) so as to comply with the inequality in [Proposition 2.2.2\(ii\)](#), starting from an initial value  $\gamma_0$  and decreasing it when necessary. The next result shows that  $\gamma_0$  is decremented only a finite number of times along the iterations, and therefore  $\gamma_k$  is positive and eventually constant. In the rest of the chapter we will denote  $\gamma_\infty$  such asymptotic value of  $\gamma_k$ .

**Lemma 2.3.1.** *Let  $(\gamma_k)_{k \in \mathbb{N}}$  be the sequence of stepsize parameters computed by [Algorithm 1](#), and let  $\gamma_\infty = \min_{i \in \mathbb{N}} \gamma_i$ . Then for all  $k \in \mathbb{N}$ ,*

$$\gamma_k \geq \gamma_\infty \geq \min \{ \gamma_0, \sigma(1 - \beta)/L_f \} > 0.$$

*Proof.* Let  $(\gamma_k)_{k \in \mathbb{N}}$  be the sequence of stepsize parameters computed by [Algorithm 1](#). To arrive to a contradiction, suppose that  $k_0$  is the smallest element of  $\mathbb{N}$  such that

$$\gamma_{k_0} < \min \{ \gamma_0, \sigma(1 - \beta)/L_f \}.$$

Clearly,  $k_0 \geq 1$ . Moreover  $\sigma^{-1}\gamma_{k_0}$  must satisfy the condition in [step 3](#): for some  $w \in \mathbb{R}^n$  (corresponding to  $w^k = x^k + \tau_k d^k$  selected before going back to [step 1](#) after the condition in [step 3](#) is passed, which might differ from the final value of  $w^k$  after [step 3](#) is passed)

$$\varphi(T_{\sigma^{-1}\gamma_{k_0}}(w)) > \varphi_{\sigma^{-1}\gamma_{k_0}}(w) - \frac{\beta\sigma^{-1}\gamma_{k_0}}{2} \|R_{\sigma^{-1}\gamma_{k_0}}(w)\|^2.$$

But from [Proposition 2.2.2\(ii\)](#) we also have

$$\begin{aligned}\varphi(T_{\sigma^{-1}\gamma_{k_0}}(w)) &\leq \varphi_{\sigma^{-1}\gamma_{k_0}}(w) - \frac{\sigma^{-1}\gamma_{k_0}}{2}(1 - \sigma^{-1}\gamma_{k_0}L_f)\|R_{\sigma^{-1}\gamma_{k_0}}(w)\|^2 \\ &\leq \varphi_{\sigma^{-1}\gamma_{k_0}}(w) - \frac{\beta\sigma^{-1}\gamma_{k_0}}{2}\|R_{\sigma^{-1}\gamma_{k_0}}(w)\|^2,\end{aligned}$$

where last inequality follows from  $\sigma^{-1}\gamma_{k_0} < (1 - \beta)/L_f$ . This leads to a contradiction, therefore  $\gamma_k \geq \min\{\gamma_0, \sigma(1 - \beta)/L_f\}$  as claimed. That  $\gamma_k$  is asymptotically constant follows since the sequence  $(\gamma_k)_{k \in \mathbb{N}}$  is non-increasing.  $\square$

**Remark 2.3.2.** In [Algorithm 1](#):

- (i) Selecting  $\beta = 0$  and  $d^k \equiv 0$ ,  $\tau_k \equiv 0$  for all  $k$  yields the classical forward-backward splitting with backtracking on  $\gamma$  [\[10, §3\]](#).
- (ii) Substituting step 4 with  $x^{k+1} \leftarrow w^k$  yields a classical line-search method for the problem of minimizing  $\varphi_\gamma^{\text{FB}}$ , where a suitable  $\gamma$  is adaptively determined. However, extensive numerical experience has shown that even though this variant seems to always converge, our choice  $x^{k+1} \leftarrow T_{\gamma_k}(w^k)$  usually performs better in practice, in terms of number of forward-backward steps, cf. [Section 2.5](#).
- (iii) Step 4 comes at no additional cost once  $\tau_k$  has been determined by means of a line-search. In fact, in order to evaluate  $\varphi_{\gamma_k}^{\text{FB}}(w^k)$  and test the condition in step 2, the evaluation of  $T_{\gamma_k}(w^k)$  is required.
- (iv) When  $L_f$  is known and  $\gamma_0 \in (0, (1 - \beta)/L_f]$ , the condition in step 3 never holds, see [Proposition 2.2.2\(ii\)](#). In this case MinFBE reduces to [Algorithm 2](#): without loss of generality we will focus the analysis on [Algorithm 1](#).

$\square$

---

**Algorithm 2** MinFBE with constant  $\gamma$ 

---

**Inputs:**  $x^0 \in \mathbb{R}^n$ ,  $\beta \in [0, 1)$ ,  $\gamma \in (0, (1 - \beta)/L_f]$

**Initialize:**  $k = 0$

- 1: Select  $d^k$  such that  $\langle d^k, \nabla \varphi_\gamma(x^k) \rangle \leq 0$
  - 2: Select  $\tau_k \geq 0$  and set  $w^k = x^k + \tau_k d^k$  such that  $\varphi_\gamma(w^k) \leq \varphi_\gamma(x^k)$
  - 3:  $x^{k+1} = T_\gamma(w^k)$
  - 4:  $k \leftarrow k + 1$ , go to step 1
- 

**Remark 2.3.3.** In order to compute descent directions in [Algorithm 1](#), one usually needs to evaluate  $\nabla \varphi_\gamma^{\text{FB}}$  at a sequence of points. In practice, this only requires to perform matrix-vector products with  $\nabla^2 f$ , see [\(2.11\)-\(2.12\)](#), and *not* the computation of the full Hessian. For example, if  $f(x) = \frac{1}{2} \|Ax - b\|^2$  then  $\nabla \varphi_\gamma^{\text{FB}}(x) = R_\gamma(x) - \gamma A^\top [AR_\gamma(x)]$ . For general nonlinear  $f$ , the product  $\nabla^2 f(x)v$  can be approximated numerically using finite-differences formulas which only require one additional evaluation of  $\nabla f$ . If  $f$  is analytic, then one can use a complex step [\[143\]](#) to overcome numerical cancellation problems, and compute  $\nabla^2 f(x)v$  to machine precision at the cost of one evaluation of  $\nabla f$ . Finally, automatic differentiation techniques can be used to evaluate such Hessian-vector products, that only require a small multiple of  $2n$  operations in addition to those required to evaluate  $f$ , see [\[108, §8.2\]](#).  $\square$

We denote by  $\omega(x^0)$  the set of cluster points of the sequence  $(x^k)_{k \in \mathbb{N}}$  produced by [Algorithm 1](#) started from  $x^0 \in \mathbb{R}^n$ . The following result states that [Algorithm 1](#) is a descent method both for the FBE  $\varphi_\gamma^{\text{FB}}$  and for the original function  $\varphi$ , and, as it holds for FBS, that the sequence of fixed-point residuals is square-summable if the function is lower bounded.

**Proposition 2.3.4** (Subsequential convergence). *Suppose that [Assumption 2.1](#) is satisfied. Then, the following hold for the sequences generated by [Algorithm 1](#):*

- (i)  $\varphi(x^{k+1}) \leq \varphi(x^k) - \frac{\beta\gamma_k}{2} \|R_{\gamma_k}(w^k)\|^2 - \frac{\gamma_k}{2} \|R_{\gamma_k}(x^k)\|^2$  for all  $k \in \mathbb{N}$ ;

- (ii) either  $(\|R_{\gamma_k}(x^k)\|)_{k \in \mathbb{N}}$  is square summable, or  $\varphi(x^k) \rightarrow \inf \varphi = -\infty$ , in which case  $\omega(x^0) = \emptyset$ ;
- (iii)  $\omega(x^0) \subseteq \text{zer } \partial\varphi$ , i.e., every cluster point of  $(x^k)_{k \in \mathbb{N}}$  is critical;
- (iv) if  $\beta > 0$ , then either  $(\|R_{\gamma_k}(w^k)\|)_{k \in \mathbb{N}}$  is square summable and every cluster point of  $(w^k)_{k \in \mathbb{N}}$  is critical, or  $\varphi_{\gamma_k}^{\text{FB}}(w^k) \rightarrow \inf \varphi = -\infty$  in which case  $(w^k)_{k \in \mathbb{N}}$  has no cluster points.

*Proof.* We have

$$\begin{aligned}
\varphi(x^{k+1}) &\leq \varphi_{\gamma_k}^{\text{FB}}(w^k) - \frac{\beta\gamma_k}{2} \|R_{\gamma_k}(w^k)\|^2 \\
&\leq \varphi_{\gamma_k}^{\text{FB}}(x^k) - \frac{\beta\gamma_k}{2} \|R_{\gamma_k}(w^k)\|^2 \\
&\leq \varphi(x^k) - \frac{\beta\gamma_k}{2} \|R_{\gamma_k}(w^k)\|^2 - \frac{\gamma_k}{2} \|R_{\gamma_k}(x^k)\|^2,
\end{aligned} \tag{2.21}$$

where the first inequality comes from step 3, the second from step 2 and the third from [Proposition 2.2.2\(i\)](#). This shows [2.3.4\(i\)](#). Let  $\varphi_\star = \lim_{k \rightarrow \infty} \varphi(x^k)$ , which exists since  $(\varphi(x^k))_{k \in \mathbb{N}}$  is monotone. If  $\varphi_\star = -\infty$ , clearly  $\inf \varphi = -\infty$  and  $\omega(x^0) = \emptyset$  due to properness and lower semi-continuity of  $\varphi$  and to the monotonic behavior of  $(\varphi(x^k))_{k \in \mathbb{N}}$ . Otherwise, telescoping the inequality we get

$$\frac{1}{2} \sum_{i=0}^k \gamma_i (\beta \|R_{\gamma_i}(w^i)\|^2 + \|R_{\gamma_i}(x^i)\|^2) \leq \varphi(x^0) - \varphi(x^{k+1}) \leq \varphi(x^0) - \varphi_\star \tag{2.22}$$

and since  $\gamma_k$  is uniformly lower bounded by a positive number (see [Lemma 2.3.1](#)) [2.3.4\(ii\)](#) follows, hence [2.3.4\(iii\)](#). If  $\beta > 0$ , observing that for  $k$  large enough such that  $\gamma_k \equiv \gamma_\infty$  we have

$$\varphi_{\gamma_k}^{\text{FB}}(w^{k+1}) \stackrel{\text{step 2}}{\leq} \varphi_{\gamma_k}^{\text{FB}}(x^{k+1}) \stackrel{\text{step 4}}{=} \varphi_{\gamma_k}^{\text{FB}}(T_{\gamma_k}(w^k)) \leq \varphi_{\gamma_k}^{\text{FB}}(w^k),$$

similar argumentations as those for proving [2.3.4\(ii\)](#) show [2.3.4\(iv\)](#).  $\square$

An immediate consequence is the following result concerning the convergence of the fixed-point residual.

**Theorem 2.3.5.** Suppose that [Assumption 2.1](#) is satisfied, and consider the sequences generated by [Algorithm 1](#). Then,

$$\min_{i=0\dots k} \|R_{\gamma_i}(x^i)\|^2 \leq \frac{2}{(k+1)} \frac{\varphi(x^0) - \inf \varphi}{\min \{\gamma_0, \sigma(1-\beta)/L_f\}}.$$

If  $\beta > 0$ , then for all  $k \in \mathbb{N}$  we also have

$$\min_{i=0\dots k} \|R_{\gamma_i}(w^i)\|^2 \leq \frac{2}{(k+1)} \frac{\varphi(x^0) - \inf \varphi}{\beta \min \{\gamma_0, \sigma(1-\beta)/L_f\}}.$$

*Proof.* If  $\inf \varphi = -\infty$  there is nothing to prove. Otherwise, since the sequence  $(\gamma_k)_{k \in \mathbb{N}}$  is nonincreasing, from [\(2.22\)](#) we get

$$\frac{(k+1)\gamma_k}{2} \left( \min_{i=0\dots k} \|R_{\gamma_i}(x^i)\|^2 + \beta \min_{i=0\dots k} \|R_{\gamma_i}(w^i)\|^2 \right) \leq \varphi(x^0) - \inf \varphi.$$

Rearranging the terms and invoking [Lemma 2.3.1](#) gives the result.  $\square$

We now analyze the convergence properties of [Algorithm 1](#). We first consider the case where  $f$  is convex. Then we discuss the general case under the assumption that  $\varphi$  has the Kurdyka-Łojasiewicz property: in this case  $(d^k)_{k \in \mathbb{N}}$  must be uniformly bounded with respect to  $(R_{\gamma_k}(x^k))_{k \in \mathbb{N}}$  in order to ensure convergence, see [Theorem 2.3.10](#), condition which is not required in the convex case. When the directions are selected, say, according to a quasi-Newton scheme  $d^k = -B_k^{-1} \nabla \varphi_{\gamma}^{\text{FB}}(x^k)$ , boundedness of  $(B_k^{-1})_{k \in \mathbb{N}}$  will be necessary for the sake of global convergence when the Kurdyka-Łojasiewicz property holds for  $\varphi$ . The latter is however a milder assumption with respect to usual nonconvex line-search methods where  $(B_k^{-1})_{k \in \mathbb{N}}$  is required to have bounded condition number or  $(d^k)_{k \in \mathbb{N}}$  to be *gradient-oriented* (see [\[109\]](#) and the references therein).

### 2.3.1 Convergence in the convex case

We now prove that when  $f$  is convex [Algorithm 1](#) converges to the optimal objective value with the same sublinear rate as FBS. Notice that we require convexity of  $f$  (and  $g$ ), and *not* that of  $\varphi_\gamma^{\text{FB}}$  which may fail to be convex even when  $\varphi$  is.

**Theorem 2.3.6** (Global sublinear convergence). *Suppose that [Assumptions 2.1](#) and [2.2](#) are satisfied, and that  $f$  is convex. Then, for the sequences generated by [Algorithm 1](#), either  $\varphi(x^0) - \inf \varphi \geq R^2/\gamma_0$  and*

$$\varphi(x^1) - \inf \varphi \leq \frac{R^2}{2\gamma_0}, \quad (2.23)$$

*or for any  $k \in \mathbb{N}$  it holds*

$$\varphi(x^k) - \inf \varphi \leq \frac{2R^2}{k \min \{\gamma_0, \sigma(1 - \beta)/L_f\}}. \quad (2.24)$$

*Proof.* The proof is similar to that of [106, Thm. 4]. By [Proposition 2.2.5\(iii\)](#) we know that  $\varphi_\gamma^{\text{FB}} \leq \varphi_\gamma$  for any  $\gamma > 0$ . Combining this with (2.21) we get

$$\varphi(x^{k+1}) \leq \min_{x \in \mathbb{R}^n} \left\{ \varphi(x) + \frac{1}{2\gamma_k} \|x - x^k\|^2 \right\}, \quad (2.25)$$

and in particular, for  $x_\star \in \operatorname{argmin} \varphi$ ,

$$\begin{aligned} \varphi(x^{k+1}) &\leq \min_{\alpha \in [0,1]} \left\{ \varphi(\alpha x_\star + (1 - \alpha)x^k) + \frac{\alpha^2}{2\gamma_k} \|x^k - x_\star\|^2 \right\} \\ &\leq \min_{\alpha \in [0,1]} \left\{ \varphi(x^k) - \alpha(\varphi(x^k) - \inf \varphi) + \frac{R^2}{2\gamma_k} \alpha^2 \right\}, \end{aligned}$$

where the last inequality follows by convexity of  $\varphi$ . If  $\varphi(x^0) - \inf \varphi \geq R^2/\gamma_0$ , then the optimal solution of the latter problem for  $k = 0$  is  $\alpha = 1$  and we obtain (2.23). Otherwise, the optimal solution is

$$\alpha = \frac{\gamma_k(\varphi(x^k) - \inf \varphi)}{R^2} \leq \frac{\gamma_k(\varphi(x^0) - \inf \varphi)}{R^2} \leq 1,$$

and we obtain

$$\varphi(x^{k+1}) \leq \varphi(x^k) - \frac{\gamma_k(\varphi(x^k) - \inf \varphi)^2}{2R^2}.$$

Letting  $\lambda_k = \frac{1}{\varphi(x^k) - \inf \varphi}$  the latter inequality is expressed as

$$\frac{1}{\lambda_{k+1}} \leq \frac{1}{\lambda_k} - \frac{\gamma_k}{2R^2 \lambda_{k+1}^2}.$$

Multiplying both sides by  $\lambda_k \lambda_{k+1}$  and rearranging

$$\lambda_{k+1} \geq \lambda_k + \frac{\gamma_k}{2R^2} \frac{\lambda_{k+1}}{\lambda_k} \geq \lambda_k + \frac{\gamma_k}{2R^2},$$

where the latter inequality follows from the fact that  $(\varphi(x^k))_{k \in \mathbb{N}}$  is non-increasing, cf. [Proposition 2.3.4\(i\)](#). Telescoping the inequality and using [Lemma 2.3.1](#), we obtain

$$\lambda_k \geq \lambda_0 + \frac{k \min \{\gamma_0, \sigma(1 - \beta)/L_f\}}{2R^2} \geq \frac{k \min \{\gamma_0, \sigma(1 - \beta)/L_f\}}{2R^2}.$$

Rearranging, we arrive at [\(2.24\)](#). □

In the following result we see that the convergence rate of  $(x^k)_{k \in \mathbb{N}}$  is linear when close to a strong local minimum.

**Theorem 2.3.7** (Local linear convergence). *Suppose that [Assumption 2.1](#) is satisfied. Suppose further that  $f$  is convex and that  $x_*$  is a strong (global) minimum of  $\varphi$ , i.e., there exist a neighborhood  $N$  of  $x_*$  and  $c > 0$  such that*

$$\varphi(x) - \varphi(x_*) \geq \frac{c}{2} \|x - x_*\|^2, \quad \forall x \in N. \quad (2.26)$$

*Then there is  $k_0 \geq 0$  such that  $(\varphi(x^k))_{k \geq k_0}$  and  $(\varphi_{\gamma_k}^{\text{FB}}(w^k))_{k \geq k_0}$  in [Algorithm 1](#) converge  $Q$ -linearly to  $\varphi(x_*)$  with factor  $\omega$ , where*

$$\omega \leq \max \left\{ \frac{1}{2}, 1 - \frac{c}{4} \min \{\gamma_0, \sigma(1 - \beta)/L_f\} \right\} \in \left[ \frac{1}{2}, 1 \right),$$

and  $(x^k)_{k \geq k_0}$  converges  $R$ -linearly to  $x_*$ . Moreover, if  $x_*$  is a strong (global) minimum for  $\varphi_{\gamma_\infty}$ , with  $\gamma_\infty$  as in [Lemma 2.3.1](#), then  $(\varphi(w^k))_{k \geq k_0}$  also converges  $R$ -linearly to  $x_*$ .

*Proof.* If (2.26) holds, then  $\varphi$  has bounded level sets and  $\text{zer } \partial\varphi = \{x_*\}$ . In particular,  $\omega(x^0) \neq \emptyset$  and [Proposition 2.3.4\(iii\)](#) then ensures  $x^k \rightarrow x_*$ . Therefore, there is  $k_0 \in \mathbb{N}$  such that  $x^k \in N$  for all  $k \geq k_0$ . Inequality (2.25) holds, and in particular for  $k \geq k_0$

$$\begin{aligned} \varphi(x^{k+1}) &\leq \min_{\alpha \in [0,1]} \left\{ \varphi(\alpha x_* + (1-\alpha)x^k) + \frac{\alpha^2}{2\gamma_k} \|x_* - x^k\|^2 \right\} \\ &\leq \min_{\alpha \in [0,1]} \left\{ \varphi(x^k) + \alpha \left( \frac{\alpha}{c\gamma_k} - 1 \right) (\varphi(x^k) - \inf \varphi) \right\}, \end{aligned}$$

where the second inequality follows by convexity of  $\varphi$  and (2.26). The minimum of last expression is achieved for  $\alpha = \min \{1, \frac{c}{2}\gamma_k\}$ . When  $\gamma_k < 2c^{-1}$  we have the bound

$$\varphi(x^{k+1}) - \inf \varphi \leq (1 - \frac{c}{4}\gamma_k)(\varphi(x^k) - \inf \varphi).$$

When instead  $\gamma_k \geq 2c^{-1}$  we have the bound

$$\varphi(x^{k+1}) - \inf \varphi \leq (c\gamma_k)^{-1}(\varphi(x^k) - \inf \varphi) \leq \frac{1}{2}(\varphi(x^k) - \inf \varphi).$$

Therefore  $\varphi(x^{k+1}) - \inf \varphi \leq \omega(\varphi(x^k) - \inf \varphi)$ , where

$$\begin{aligned} \omega &\leq \sup_k \max \left\{ \frac{1}{2}, 1 - \frac{c}{4}\gamma_k \right\} \\ &\leq \max \left\{ \frac{1}{2}, 1 - \frac{c}{4} \min \{ \gamma_0, \sigma(1-\beta)/L_f \} \right\} \in \left[ \frac{1}{2}, 1 \right), \end{aligned}$$

last inequality following from [Lemma 2.3.1](#). This proves the claim on the sequence  $(\varphi(x^k))_{k \geq k_0}$  and using inequality (2.21) the same holds for  $(\varphi_{\gamma_k}^{\text{FB}}(w^k))_{k \geq k_0}$ . From the error bound (2.26) we obtain that  $x^k \rightarrow x_*$   $R$ -linearly. If the same error bound holds for  $\varphi_{\gamma_\infty}$ , then also  $w^k \rightarrow x_*$   $R$ -linearly.  $\square$



The introduction of  $\gamma_\infty$  in the statement above is due to the fact that  $\gamma_k$  may vary over the iterations. However, under the assumptions of [Theorem 2.2.11](#), if  $\gamma_\infty < 1/L_f$  then the requirement of  $x_\star$  to be a strong local minimizer for  $\varphi_{\gamma_\infty}$  is superfluous, as it is already implied by strong local minimality of  $x_\star$  for  $\varphi$ .

**Corollary 2.3.8** (Global linear convergence). *Suppose that [Assumption 2.1](#) is satisfied, that  $f$  is convex and that  $\varphi$  is strongly convex (e.g. if  $f$  is strongly convex). Then, the sequences  $(\varphi(x^k))_{k \in \mathbb{N}}$  and  $(\varphi_{\gamma_k}^{\text{FB}}(w^k))_{k \in \mathbb{N}}$  generated by [Algorithm 1](#) converge  $Q$ -linearly to  $\varphi_\star$ , while  $(x^k)_{k \in \mathbb{N}}$  converges  $R$ -linearly to  $x_\star$ .*

*Proof.* In this case [Theorem 2.3.7](#) applies with  $N = \mathbb{R}^n$ ,  $c = \mu_\varphi$  (the convexity modulus of  $\varphi$ ) and  $k_0 = 0$ .  $\square$

## 2.3.2 Convergence under KL assumption

We now analyze the convergence of the iterates of [Algorithm 1](#) to a critical point under the assumption that  $\varphi$  satisfies the Kurdyka-Łojasiewicz (KL) property [[95](#), [96](#), [85](#)]. For related works exploiting this property in proving convergence of optimization algorithms such as FBS we refer the reader to [[4](#), [5](#), [6](#), [21](#), [110](#)].

**Definition 2.3.9** (KL property [[21](#), Def. 3]). *A proper lower semi-continuous function  $\varphi : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  has the Kurdyka-Łojasiewicz property (KL) at  $x_\star \in \text{dom } \partial\varphi$  if there exist  $\eta \in (0, +\infty]$ , a neighborhood  $U$  of  $x_\star$ , and a continuous concave function  $\psi : [0, \eta] \rightarrow [0, +\infty)$  such that:*

- (i)  $\psi(0) = 0$ ,
- (ii)  $\psi$  is  $C^1$  on  $(0, \eta)$ ,
- (iii)  $\psi'(s) > 0$  for all  $s \in (0, \eta)$ ,
- (iv) for every  $x \in U \cap \{x \in \mathbb{R}^n \mid \varphi(x_\star) < \varphi(x) \leq \varphi(x_\star) + \eta\}$ ,

$$\psi'(\varphi(x) - \varphi(x_\star)) \text{dist}(0, \partial\varphi(x)) \geq 1.$$

We say that  $\varphi$  has the KL property on  $S \subseteq \mathbb{R}^n$  if it has the KL property on every  $x \in S$ .

Function  $\psi$  in the previous definition is usually called *desingularizing function*. All subanalytic functions which are continuous over their domain have the KL property [20]. Under the KL assumption we are able to prove the following convergence result. Once again, we remark that such property is required on the original function  $\varphi$ , rather than on the surrogate  $\varphi_\gamma^{\text{FB}}$ .

**Theorem 2.3.10.** *Suppose that [Assumptions 2.1](#) and [2.2](#) are satisfied, and that  $\varphi$  satisfies the KL property on  $\omega(x^0)$  (e.g. if it has it on  $\text{zer } \partial\varphi$ ). Suppose further that in [Algorithm 1](#)  $\beta > 0$ , and that there exist  $\bar{\tau}, c > 0$  such that  $\tau_k \leq \bar{\tau}$  and  $\|d^k\| \leq c\|R_{\gamma_k}(x^k)\|$  for all  $k \in \mathbb{N}$ . Then, the sequence of iterates  $(x^k)_{k \in \mathbb{N}}$  is either finite and ends with  $R_{\gamma_k}(x^k) = 0$ , or converges to a critical point  $x_\star$  of  $\varphi$ .*

*Proof.* The case where the sequence is finite does not deserve any further investigation, therefore we assume that  $(x^k)_{k \in \mathbb{N}}$  is infinite. We then assume that  $R_{\gamma_k}(x^k) \neq 0$  which implies through [Proposition 2.3.4](#) that  $\varphi(x^{k+1}) < \varphi(x^k)$ . Due to (2.45), the KL property for  $\varphi$ , and [Lemma 2.A.7](#), there exist  $\varepsilon, \eta > 0$  and a continuous concave function  $\psi : [0, \eta] \rightarrow [0, +\infty)$  such that for all  $x$  with  $\text{dist}_{\omega(x^0)}(x) < \varepsilon$  and  $\varphi(x^\star) < \varphi(x) < \varphi(x_\star) + \eta$  one has

$$\psi'(\varphi(x) - \varphi(x_\star)) \text{dist}(0, \partial\varphi(x)) \geq 1.$$

According to [Proposition 2.A.6](#) there is  $k_1 \in \mathbb{N}$  such that  $\text{dist}_{\omega(x^0)}(x^k) < \varepsilon$  for all  $k \geq k_1$ . Furthermore, since  $\varphi(x^k)$  converges to  $\varphi(x_\star)$  there exists a  $k_2$  such that  $\varphi(x^k) < \varphi(x_\star) + \eta$  for all  $k \geq k_2$ . Take  $\bar{k} = \max\{k_1, k_2\}$ . Then for every  $k \geq \bar{k}$  we have

$$\psi'(\varphi(x^k) - \varphi(x_\star)) \text{dist}(0, \partial\varphi(x^k)) \geq 1.$$

From [Proposition 2.3.4\(i\)](#)

$$\varphi(x^{k+1}) \leq \varphi(x^k) - \frac{\beta\gamma_k}{2} \|R_{\gamma_k}(w^k)\|^2.$$

For every  $k > 0$  let  $\tilde{\nabla}\varphi(x^k) = \nabla f(x^k) - \nabla f(w^{k-1}) + R_{\gamma_{k-1}}(w^{k-1})$ . Since  $R_{\gamma_{k-1}}(w^{k-1}) \in \nabla f(w^{k-1}) + \partial g(x^k)$ , then  $\tilde{\nabla}\varphi(x^k) \in \partial\varphi(x^k)$  and

$$\begin{aligned} \|\tilde{\nabla}\varphi(x^k)\| &\leq \|\nabla f(x^k) - \nabla f(w^{k-1})\| + \|R_{\gamma_{k-1}}(w^{k-1})\| \\ &= (1 + \gamma_{k-1}L_f) \|R_{\gamma_{k-1}}(w^{k-1})\|. \end{aligned}$$

From [\(2.46\)](#)

$$\psi'(\varphi(x^k) - \varphi(x_*)) \geq \frac{1}{\|\tilde{\nabla}\varphi(x^k)\|} \geq \frac{1}{(1 + \gamma_{k-1}L_f) \|R_{\gamma_{k-1}}(w^{k-1})\|}.$$

Let  $\Delta_k = \psi(\varphi(x^k) - \varphi(x_*))$ . By concavity of  $\psi$  and [Proposition 2.3.4\(i\)](#)

$$\begin{aligned} \Delta_k - \Delta_{k+1} &\geq \psi'(\varphi(x^k) - \varphi(x_*))(\varphi(x^k) - \varphi(x^{k+1})) \\ &\geq \frac{\beta\gamma_k}{2(1 + \gamma_{k-1}L_f)} \frac{\|R_{\gamma_k}(w^k)\|^2}{\|R_{\gamma_{k-1}}(w^{k-1})\|} \\ &\geq \frac{\beta\gamma_{\min}}{2(1 + \gamma_0L_f)} \frac{\|R_{\gamma_k}(w^k)\|^2}{\|R_{\gamma_{k-1}}(w^{k-1})\|} \end{aligned}$$

where  $\gamma_{\min} = \min\{\gamma_0, \sigma(1 - \beta)/L_f\}$ , see [Lemma 2.3.1](#), or

$$\|R_{\gamma_k}(w^k)\|^2 \leq \alpha(\Delta_k - \Delta_{k+1}) \|R_{\gamma_{k-1}}(w^{k-1})\| \quad (2.27)$$

where  $\alpha = 2(1 + \gamma_0L_f)/(\beta\gamma_{\min})$ . Applying [Lemma 2.A.5](#) with

$$\delta_k = \alpha\Delta_k, \quad \beta_k = \|R_{\gamma_{k-1}}(w^{k-1})\|,$$

we conclude that  $\sum_{k=0}^{\infty} \|R_{\gamma_k}(w^k)\| < \infty$ . From [\(2.42\)](#), using the fact that

$\gamma_k \leq \gamma_0$  for all  $k$ , then it follows that

$$\sum_{k=0}^{\infty} \|x^{k+1} - x^k\| < \infty.$$

Then  $(x^k)_{k \in \mathbb{N}}$  is a Cauchy sequence, hence it converges to a point that, by [Proposition 2.3.4](#), is a critical point  $x_*$  of  $\varphi$ .  $\square$

In case where  $\varphi$  is subanalytic, the desingularizing function can be taken of the form  $\psi(s) = \sigma s^{1-\theta}$ , for  $\sigma > 0$  and  $\theta \in [0, 1)$  [\[20\]](#). In this case, the condition in [Definition 2.3.9\(iv\)](#) is referred to as Łojasiewicz inequality. Depending on the value of  $\theta$  we can derive local convergence rates for [Algorithm 1](#).

**Theorem 2.3.11** (Local linear convergence). *Suppose that [Assumptions 2.1](#) and [2.2](#) are satisfied, and that  $\varphi$  satisfies the KL property on  $\omega(x^0)$  (e.g. if it has it on  $\text{zer } \partial\varphi$ ) with*

$$\psi(s) = \sigma s^{1-\theta} \quad \text{for some } \sigma > 0 \text{ and } \theta \in (0, \tfrac{1}{2}]. \quad (2.28)$$

*Suppose further that in [Algorithm 1](#)  $\beta > 0$ , and that there exist  $\bar{\tau}, c > 0$  such that  $\tau_k \leq \bar{\tau}$  and  $\|d^k\| \leq c\|R_{\gamma_k}(x^k)\|$  for all  $k \in \mathbb{N}$ . Then, the sequence of iterates  $(x^k)_{k \in \mathbb{N}}$  converges to a point  $x_* \in \text{zer } \partial\varphi$  with  $R$ -linear rate.*

*Proof.* [Theorem 2.3.10](#) ensures that  $(x^k)_{k \in \mathbb{N}}$  converges to a critical point, be it  $x_*$ . We know from [Lemma 2.3.1](#) that eventually  $\gamma_k = \gamma_\infty > 0$ , therefore we assume  $k$  is large enough for this purpose and indicate  $\gamma$  in place of  $\gamma_k$  for simplicity. Denoting  $A_k = \sum_{i=k}^{\infty} \|x^{i+1} - x^i\|$  clearly  $A_k \geq \|x^k - x_*\|$ , so we will prove that  $A_k$  converges linearly to zero to obtain the result. Note that by [\(2.42\)](#) we know that

$$\|x^{i+1} - x^i\| \leq \gamma \|R_\gamma(w^i)\| + \bar{\tau}c(1 + \gamma L_f) \|R_\gamma(w^{i-1})\|.$$

Therefore we can upper bound  $A_k$  as follows

$$\begin{aligned} A_k &\leq \bar{\tau}c(1 + \gamma L_f) \|R_\gamma(w^{k-1})\| + (\gamma + \bar{\tau}c(1 + \gamma L_f)) \sum_{i=k}^{\infty} \|R_\gamma(w^i)\| \\ &\leq (\gamma + \bar{\tau}c(1 + \gamma L_f)) \sum_{i=k-1}^{\infty} \|R_\gamma(w^i)\|, \end{aligned} \quad (2.29)$$

and reduce the problem to proving linear convergence of the sequence of  $B_k = \sum_{i=k}^{\infty} \|R_\gamma(w^i)\|$ . When  $\psi$  is as in (2.28), for sufficiently large  $k$  the KL inequality reads

$$\varphi(x^k) - \varphi(x_\star) \leq [\sigma(1 - \theta) \|v^k\|]^{\frac{1}{\theta}}, \quad \forall v^k \in \partial\varphi(x^k).$$

Taking  $v^k = \nabla f(x^k) - \nabla f(w^{k-1}) + R_\gamma(w^{k-1}) \in \partial\varphi(x^k)$ , this in turn yields

$$\varphi(x^k) - \varphi(x_\star) \leq [\sigma(1 - \theta)(1 + \gamma L_f) \|R_\gamma(w^{k-1})\|]^{\frac{1}{\theta}}, \quad (2.30)$$

(see the proof of [Theorem 2.3.10](#)). Inequality (2.27) holds, for sufficiently large  $k$ , with  $\Delta_k = \sigma(\varphi(x^k) - \varphi(x_\star))^{1-\theta}$  in this case. Applying [Lemma 2.A.5](#) with

$$\delta_k = \alpha \Delta_k, \quad \beta_k = \|R_\gamma(w^{k-1})\| = B_{k-1} - B_k,$$

we obtain

$$\begin{aligned} B_k &\leq (B_{k-1} - B_k) + \sigma(\varphi(x^k) - \varphi(x_\star))^{1-\theta} \\ &\leq (B_{k-1} - B_k) + \sigma [\sigma(1 - \theta)(1 + \gamma L_f)(B_{k-1} - B_k)]^{\frac{1-\theta}{\theta}}, \end{aligned}$$

where the second inequality is due to (2.30). Since  $B_{k-1} - B_k \rightarrow 0$ , then for  $k$  large enough it holds that  $\sigma(1 + \gamma L_f)(B_{k-1} - B_k) \leq 1$ , and the last term in the previous chain of inequalities is increasing in  $\theta$  when  $\theta \in (0, \frac{1}{2}]$ . Therefore  $B_k$  eventually satisfies

$$B_k \leq C(B_{k-1} - B_k),$$

where  $C > 0$ , and so  $B_k \leq [C/(1+C)]B_{k-1}$ , i.e.,  $B_k$  converges to zero  $Q$ -linearly. This in turn implies that  $\|x^k - x_\star\|$  converges to zero with  $R$ -linear rate. Furthermore,

$$\begin{aligned}
\|w^k - x_\star\| &= \|x^k - x_\star + \tau_k d^k\| \\
&\leq \|x^k - x_\star\| + \bar{\tau} c \|R_{\gamma_k}(x^k)\| \\
&= \|x^k - x_\star\| + \bar{\tau} c \gamma_k^{-1} \|T_{\gamma_k}(x^k) - x^k\| \\
&\leq (1 + \bar{\tau} c \gamma_k^{-1}) \|x^k - x_\star\| + \bar{\tau} c \gamma_k^{-1} \|T_{\gamma_k}(x^k) - T_{\gamma_k}(x_\star)\| \\
&\leq (1 + \bar{\tau} c \gamma_k^{-1}) \|x^k - x_\star\| \\
&\quad + \bar{\tau} c \gamma_k^{-1} \|x^k - \gamma_k \nabla f(x^k) - x_\star + \gamma_k \nabla f(x_\star)\| \\
&\leq (1 + \bar{\tau} c (2\gamma_k^{-1} + L_f)) \|x^k - x_\star\|,
\end{aligned}$$

where the last two inequalities follow by nonexpansiveness of  $\text{prox}_{\gamma g}$  and Lipschitz continuity of  $\nabla f$ . Since  $\gamma_k$  is lower bounded by a positive quantity, then we deduce that also  $\|w^k - x_\star\|$  converges  $R$ -linearly to zero.  $\square$

## 2.4 Quasi-Newton methods

We now turn our attention to choices of the direction  $d^k$  in [Algorithm 1](#). Applying classical quasi-Newton methods [\[48\]](#) to the problem of minimizing  $\varphi_\gamma^{\text{FB}}$  yields, starting from a given  $x^0$ ,

$$\begin{aligned}
d^k &= -B_k^{-1} \nabla \varphi_\gamma^{\text{FB}}(x^k), \\
x^{k+1} &= x^k + \tau_k d^k,
\end{aligned}$$

where  $B_k$  is nonsingular and chosen so as to approximate (in some sense) the Hessian of  $\varphi_\gamma^{\text{FB}}$  at  $x^k$ , and stepsize  $\tau_k > 0$  is selected with a line-search procedure enforcing a sufficient decrease condition. However, the convergence properties of quasi-Newton methods are quite restrictive.

The BFGS algorithm is guaranteed to converge, in the sense that

$$\liminf_{k \rightarrow \infty} \|\nabla \varphi_{\gamma}^{\text{FB}}(x^k)\| = 0,$$

when the objective is convex [123]. Its limited memory variant, L-BFGS, requires strong convexity to guarantee convergence, and in that case the cost is shown to converge  $R$ -linearly to the optimal value [93]. Moreover, there exist examples of nonconvex function for which quasi-Newton methods need not converge to critical points [39, 98, 99, 40].

To overcome this, we consider quasi-Newton directions in the setting of Algorithm 1. The resulting methods enjoy the same global convergence properties illustrated in Section 2.3 and superlinear asymptotic convergence under standard assumptions: we will assume, as it is usual, (strict) differentiability of  $\nabla \varphi_{\gamma}^{\text{FB}}$  and nonsingularity of  $\nabla^2 \varphi_{\gamma}^{\text{FB}}$  at a critical point. Properties of  $f$  and  $g$  that guarantee these requirements were discussed in Theorems 2.2.11 and 2.2.10: if  $\gamma = \gamma_{\infty}$  is as in Lemma 2.3.1, then (strict) differentiability of  $\nabla \varphi_{\gamma}^{\text{FB}}$  at  $x_{\star} \in \text{zer } \partial \varphi$  and positive definiteness of  $\nabla^2 \varphi_{\gamma}^{\text{FB}}(x_{\star})$  are ensured if Assumption 2.4 (2.4<sup>+</sup>) holds,  $x_{\star}$  is a strong local minimum for  $\varphi$ , and  $\gamma < 1/L_f$ .

The following result gives for the proposed algorithmic scheme the analogous of the Dennis-Moré condition, see [47, Thm. 2.2] and [84, Thm. 3.3]. Differently from the cited results, we fit the analysis to our algorithmic framework where an additional forward-backward step is operated. Furthermore, in Theorem 2.4.2 we will see how achieving superlinear convergence is possible without the need to ensure *sufficient* decrease in the objective, or even to consider direction of strict descent, but simply with the nonincrease conditions of steps 1 and 2. This contrasts with the usual requirements of classical line-search methods, where instead a sufficient decrease must be enforced in order for the sequence of iterates to converge. In Algorithm 1, in fact, such decrease is guaranteed by the final update in step 4.

**Theorem 2.4.1.** *Suppose that Assumption 2.1 is satisfied, and let  $\gamma > 0$ . Sup-*

pose that  $\nabla\varphi_\gamma^{\text{FB}}$  is strictly differentiable at  $x_\star$ , and that  $\nabla^2\varphi_\gamma^{\text{FB}}(x_\star)$  is nonsingular. Let  $(B_k)_{k \in \mathbb{N}}$  be a sequence of nonsingular  $\mathbb{R}^{n \times n}$ -matrices and suppose that for some  $x^0 \in \mathbb{R}^n$  the sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  generated by

$$w^k = x^k - B_k^{-1} \nabla\varphi_\gamma^{\text{FB}}(x^k) \quad \text{and} \quad x^{k+1} = T_\gamma(w^k)$$

converge to  $x_\star$ . If  $x^k, w^k \notin \text{zer } \partial\varphi$  for all  $k \geq 0$  and

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - \nabla^2\varphi_\gamma^{\text{FB}}(x_\star))(w^k - x^k)\|}{\|w^k - x^k\|} = 0, \quad (2.31)$$

then  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge  $Q$ -superlinearly to  $x_\star$ .

*Proof.* Since  $w^k = x^k - B_k^{-1} \nabla\varphi_\gamma^{\text{FB}}(x^k)$ , letting  $k \rightarrow \infty$  and using (2.31) we have that

$$\begin{aligned} 0 &\leftarrow - \frac{(B_k - \nabla^2\varphi_\gamma^{\text{FB}}(x_\star))(w^k - x^k)}{\|w^k - x^k\|} \\ &= \frac{\nabla\varphi_\gamma^{\text{FB}}(x^k) + \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)(w^k - x^k)}{\|w^k - x^k\|} \\ &= \frac{\nabla\varphi_\gamma^{\text{FB}}(x^k) - \nabla\varphi_\gamma^{\text{FB}}(w^k) + \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)(w^k - x^k)}{\|w^k - x^k\|} + \frac{\nabla\varphi_\gamma^{\text{FB}}(w^k)}{\|w^k - x^k\|}. \end{aligned}$$

By strict differentiability of  $\nabla\varphi_\gamma^{\text{FB}}$  at  $x_\star$  we obtain

$$\lim_{k \rightarrow \infty} \frac{\|\nabla\varphi_\gamma^{\text{FB}}(w^k)\|}{\|w^k - x^k\|} = 0 \quad (2.32)$$

By nonsingularity of  $\nabla^2\varphi_\gamma^{\text{FB}}(x_\star)$  and since  $w^k \rightarrow x_\star$ , there exist  $\alpha > 0$  such that  $\|\nabla\varphi_\gamma^{\text{FB}}(x^k)\| \geq \alpha\|x^k - x_\star\|$  for  $k$  large enough. Therefore, for  $k$  sufficiently large,

$$\frac{\|\nabla\varphi_\gamma^{\text{FB}}(w^k)\|}{\|w^k - x^k\|} \geq \frac{\alpha\|w^k - x_\star\|}{\|w^k - x^k\|} \geq \frac{\alpha\|w^k - x_\star\|}{\|w^k - x_\star\| + \|x^k - x_\star\|}.$$



Using (2.32) we get

$$\lim_{k \rightarrow \infty} \frac{\|w^k - x_\star\|}{\|w^k - x_\star\| + \|x^k - x_\star\|} = \lim_{k \rightarrow \infty} \frac{\|w^k - x_\star\|/\|x^k - x_\star\|}{\|w^k - x_\star\|/\|x^k - x_\star\| + 1} = 0,$$

from which we obtain

$$\lim_{k \rightarrow \infty} \frac{\|w^k - x_\star\|}{\|x^k - x_\star\|} = 0. \quad (2.33)$$

Finally,

$$\begin{aligned} \|x^{k+1} - x_\star\| &= \|T_\gamma(w^k) - T_\gamma(x_\star)\| \\ &= \|\text{prox}_{\gamma g}(w^k - \gamma \nabla f(w^k)) - \text{prox}_{\gamma g}(x_\star - \gamma \nabla f(x_\star))\| \\ &\leq \|w^k - \gamma \nabla f(w^k) - x_\star + \gamma \nabla f(x_\star)\| \\ &\leq (1 + \gamma L_f) \|w^k - x_\star\|, \end{aligned} \quad (2.34)$$

where the first inequality follows from nonexpansiveness of  $\text{prox}_{\gamma g}$  and the second from Lipschitz continuity of  $\nabla f$ . Using (2.34) in (2.33) we obtain that  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge  $Q$ -superlinearly to  $x_\star$ .  $\square$

To obtain superlinear convergence of [Algorithm 1](#) when quasi-Newton directions are used and condition (2.31) on the sequence  $(B_k)_{k \in \mathbb{N}}$  holds, we must verify that eventually  $\varphi_\gamma^{\text{FB}}(x^k + d^k) \leq \varphi_\gamma^{\text{FB}}(x^k)$ , so that the step-size  $\tau_k = 1$  is accepted in step 2 and the iterations reduce to those described in [Theorem 2.4.1](#).

**Theorem 2.4.2.** *Suppose that [Assumption 2.1](#) is satisfied, and that in [Algorithm 1](#) direction  $d^k$  is set as*

$$d^k = -B_k^{-1} \nabla \varphi_{\gamma_k}^{\text{FB}}(x^k)$$

*for a sequence of nonsingular matrices  $(B_k)_{k \in \mathbb{N}}$  satisfying (2.31), with  $\tau_k = 1$  being tried first in step 2. Let  $\gamma = \gamma_\infty$  as in [Lemma 2.3.1](#), and suppose further that the sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge to a critical point  $x_\star$  at which*

$\nabla\varphi_\gamma^{\text{FB}}$  is continuously semidifferentiable with  $\nabla^2\varphi_\gamma^{\text{FB}}(x_\star) \succ 0$ . Then,  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge  $Q$ -superlinearly to  $x_\star$ .

*Proof.* From [Proposition 1.A.2\(a\)](#) it follows that  $\nabla\varphi_\gamma^{\text{FB}}$  is strictly differentiable and continuously semidifferentiable at  $x_\star$ . Moreover, we know from [Lemma 2.3.1](#) that eventually  $\gamma_k = \gamma_\infty > 0$ . Therefore we assume that  $k$  is large enough for this purpose and indicate  $\gamma$  in place of  $\gamma_k$  for simplicity. We denote for short  $g^k = \nabla\varphi_\gamma^{\text{FB}}(x^k)$ . In [Algorithm 1](#)

$$w^k - x^k = \tau_k d^k = -\tau_k B_k^{-1} g^k,$$

and by [\(2.31\)](#) and Cauchy-Schwarz inequality

$$\begin{aligned} \frac{\|(B_k - \nabla^2\varphi_\gamma^{\text{FB}}(x_\star))(w^k - x^k)\|}{\|w^k - x^k\|} &= \frac{\|g^k + \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)d^k\|}{\|d^k\|} \\ &\geq \left| \frac{\langle d^k, g^k + \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)d^k \rangle}{\|d^k\|^2} \right| \rightarrow 0. \end{aligned}$$

Therefore

$$-\langle g^k, d^k \rangle = \langle d^k, \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)d^k \rangle + o(\|d^k\|^2). \quad (2.35)$$

Since  $\nabla^2\varphi_\gamma^{\text{FB}}(x_\star)$  is positive definite, then there is  $\eta > 0$  such that for sufficiently large  $k$

$$-\langle g^k, d^k \rangle \geq \eta\|d^k\|^2. \quad (2.36)$$

Since  $D\nabla\varphi_\gamma^{\text{FB}}$  is continuous at  $x_\star$  and  $x^k \rightarrow x_\star$ , we have

$$\|D\nabla\varphi_\gamma^{\text{FB}}(x^k)[d^k] - \nabla^2\varphi_\gamma^{\text{FB}}(x_\star)d^k\| = o(\|d^k\|). \quad (2.37)$$

Next, since  $x^k \rightarrow x_\star$ , for  $k$  large enough  $\nabla\varphi_\gamma^{\text{FB}}$  is semidifferentiable at  $x^k$

and we can expand  $\varphi_\gamma^{\text{FB}}$  around  $x^k$  using [133, Ex. 13.7(c)] to obtain

$$\begin{aligned}\varphi_\gamma^{\text{FB}}(x^k + d^k) - \varphi_\gamma^{\text{FB}}(x^k) &= \langle g^k, d^k \rangle + \frac{1}{2} \langle d^k, D\nabla \varphi_\gamma^{\text{FB}}(x^k)[d^k] \rangle + o(\|d^k\|^2) \\ &= \langle g^k, d^k \rangle + \frac{1}{2} \langle d^k, \nabla^2 \varphi_\gamma^{\text{FB}}(x_\star) d^k \rangle + o(\|d^k\|^2) \\ &= \frac{1}{2} \langle g^k, d^k \rangle + o(\|d^k\|^2),\end{aligned}$$

where the second equality is due to (2.37), and the last equality is due to (2.35). Therefore, using (2.36), for sufficiently large  $k$

$$\varphi_\gamma^{\text{FB}}(x^k + d^k) - \varphi_\gamma^{\text{FB}}(x^k) \leq -\frac{\eta}{2} \|d^k\|^2 < 0.$$

i.e.,  $\tau_k = 1$  satisfies the non-increase condition. As a consequence, Algorithm 1 eventually reduces to the iterations of Theorem 2.4.1 and the proof follows.  $\square$

## 2.4.1 BFGS

The sequence  $(B_k)_{k \in \mathbb{N}}$  can be computed using BFGS updates: starting from  $B_0 \succ 0$ , use vectors

$$s^k = w^k - x^k, \quad y^k = \nabla \varphi_\gamma(w^k) - \nabla \varphi_\gamma(x^k), \quad (2.38a)$$

to compute

$$B_{k+1} = \begin{cases} B_k + \frac{y^k (y^k)^\top}{\langle y^k, s^k \rangle} - \frac{B_k s^k (B_k s^k)^\top}{\langle s^k, B_k s^k \rangle} & \text{if } \langle s^k, y^k \rangle > 0, \\ B_k & \text{otherwise.} \end{cases} \quad (2.38b)$$

Note that in this way  $B_k \succ 0$ , for all  $k \geq 0$ , and  $d^k = -B^{-1} \nabla \varphi_\gamma(x^k)$  is always a direction of descent for  $\varphi_\gamma^{\text{FB}}$ . No matrix inversion is needed to compute  $d^k$  in practice, since it is possible to perform the inverse updates of (2.38b) directly producing the sequence  $(B_k^{-1})_{k \in \mathbb{N}}$ , see [48, 108].

In light of the convergence results for MinFBE given in Section 2.3 we

prove superlinear convergence for the update (2.38) under the following assumption.

**Assumption 2.5.** *Either of the following holds:*

- (i)  $\varphi$  is convex and has a strong local minimum  $x_*$ ;
- (ii)  $\varphi$  has the KL property on  $\omega(x^0)$  with  $\psi(s) = \sigma s^{1-\theta}$ , where  $\sigma > 0$  and  $\theta \in (0, \frac{1}{2}]$ , and in MinFBE the stepsize  $\tau_k$  is bounded and there is  $c > 0$  such that for all  $k \in \mathbb{N}$ ,  $\|d^k\| \leq c \|\nabla \varphi_{\gamma_k}(x^k)\|$ .

**Theorem 2.4.3.** *Suppose that Assumptions 2.1 and 2.5 hold, and that directions  $d^k$  in MinFBE are set as*

$$d^k = -B_k^{-1} \nabla \varphi_{\gamma_k}(x^k) \quad \text{with } B_k \text{ as in (2.38),}$$

and with  $\tau_k = 1$  being tried first in step 2. Let  $\gamma = \gamma_\infty$  as in Lemma 2.3.1, and suppose further that the sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge to a critical point  $x_*$  at which  $\nabla \varphi_\gamma$  is calmly semidifferentiable (see Proposition 2.A.3) with  $\nabla^2 \varphi_\gamma(x_*) \succ 0$ . Then,  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge  $Q$ -superlinearly to  $x_*$ .

*Proof.* Suppose that Assumption 2.5(i) holds. Since  $x_* \in \text{zer } \partial \varphi$  and  $\nabla^2 \varphi_\gamma(x_*) \succ 0$ , it follows that  $x_*$  is a strong local minimizer of  $\varphi_\gamma$ , hence of  $\varphi$  in light of Propositions 2.2.2(i) and 2.2.3(i). Theorem 2.3.7 then ensures that  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge linearly to  $x_*$ . If instead Assumption 2.5(ii) holds, then we can invoke Theorem 2.3.11 (since it holds  $\|\nabla \varphi_{\gamma_k}(x^k)\| \leq (1 + \gamma_0 L_f) \|R_{\gamma_k}(x^k)\|$ ) to infer that  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  converge linearly to a critical point, be it  $x_*$ . In both cases we can apply Proposition 2.A.3 and for  $k$  sufficiently large

$$\frac{\|y^k - \nabla^2 \varphi_\gamma(x_*) s^k\|}{\|s^k\|} \leq L \max \{\|w^k - x_*\|, \|x^k - x_*\|\}. \quad (2.39)$$

Since the convergence is linear, then the right-hand side of (2.39) is summable. With similar arguments to those of [47, Lem. 3.2] we can see that eventually  $\langle s^k, y^k \rangle > 0$ . Therefore we can apply [30, Thm. 3.2], which

ensures that condition (2.31) holds. The result follows then from [Theorem 2.4.2](#).  $\square$

### 2.4.2 L-BFGS

When dealing with a large number of variables, storing (and updating) approximations of the Hessian matrix (or its inverse) may be impractical. Limited-memory quasi-Newton methods remedy this by storing, instead of a dense  $n \times n$  matrix, only a few most recent pairs  $(s^k, y^k)$  implicitly representing such approximation. The limited-memory BFGS method (L-BFGS) is probably the most widely used method of this class, and was first introduced in [93]. It is based on the BFGS update, but uses at iteration  $k$  only the most recent  $\tilde{m} = \min \{m, k\}$  pairs (here  $m$  is a parameter, usually  $m \in \{3, \dots, 20\}$ ) to compute a descent direction:  $d^k$  is obtained using a procedure known as *two-loop recursion* [107], so that no matrix storage is required, and in fact only  $O(n)$  operations are needed. For this reason L-BFGS is better suited for large scale applications. Similarly to BFGS, a safeguard is used to make sure that  $\langle s^k, y^k \rangle > 0$ , so that  $d^k$  is always a descent direction for  $\varphi_{\gamma_k}^{\text{FB}}$ .

**Remark 2.4.4.** In both BFGS and L-BFGS, the condition  $\langle s^k, y^k \rangle > 0$  is sufficient to ensure the positive definiteness of the Hessian approximation, hence the fact that  $d^k$  is a descent direction. Therefore, in [Algorithm 1](#) one can simply check such condition and discard the update when it does not hold. Other methods were proposed in the literature to ensure convergence of quasi-Newton methods in the nonconvex case, by Powell (see [108, §18.3]) and Li, Fukushima [88]. In our experience, no significant advantage is gained when using these techniques in [Algorithm 1](#). Moreover, no such care is required for [Algorithm 1](#) to converge to a critical point, and under the assumptions of [Theorem 2.4.2](#) the condition  $\langle s^k, y^k \rangle > 0$  will eventually always hold (see the proof of [Theorem 2.4.2](#) for details).  $\square$

## 2.5 Simulations

We now present numerical results obtained with the proposed method. In all the results, we indicate in parenthesis the choice of directions for MinFBE. We set  $\beta = 0.05$  in [Algorithm 1](#), therefore if  $L_f$  is known then we set a constant  $\gamma = 0.95/L_f$ . To determine the stepsize  $\tau_k$  in [Algorithm 1](#) we use backtracking, starting with  $\tau_k = 1$  and reducing it until  $\varphi_{\gamma_k}^{\text{FB}}(x^k + \tau_k d^k) \leq \varphi_{\gamma_k}^{\text{FB}}(x^k)$  holds.

Among the other algorithms, for each choice of descent directions we also compare [Algorithm 1](#) with the corresponding classical line-search method, see [Remark 2.3.2\(ii\)](#). In this case we use a line-search procedure, inspired by [\[80, §II.3.3\]](#), enforcing the usual Wolfe conditions: although simpler, in our tests this strategy performed favorably with respect to other algorithms, see for example [\[16, §1.2\]](#), [\[108, §3\]](#), [\[63, §2.6\]](#). We always set the memory parameter  $m = 5$  when computing L-BFGS directions. All experiments were performed in MATLAB, and the implementation of the methods used in the tests are available.<sup>1</sup>

### 2.5.1 Lasso

The problem is to find a sparse representation of a vector  $b \in \mathbb{R}^m$  as combination of the columns of  $A \in \mathbb{R}^{m \times n}$ . This is done by minimizing  $\varphi = f + g$  where

$$f(x) = \frac{1}{2} \|Ax - b\|_2^2, \quad g(x) = \lambda \|x\|_1.$$

The proximal mapping of  $g$  is the soft-thresholding operation, while the computationally relevant operation here is the evaluation of  $f$  and  $\nabla f$ , which involves matrix-vector products with  $A$  and  $A^\top$ . Parameter  $\lambda$  modulates between a small least squares residual and a sparse solution vector  $x_*$ , *i.e.*, the larger the  $\lambda$  the more zero coefficients  $x_*$  has. In particular,  $\lambda_{\max} = \|A^\top b\|_\infty$  is the minimum value such that for  $\lambda \geq \lambda_{\max}$

---

<sup>1</sup><http://kul-forbes.github.io/ForBES/>

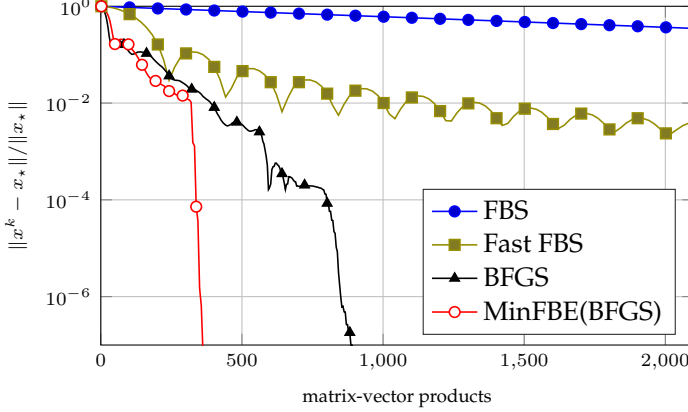
the solution is  $x_\star = 0$ . We have  $L_f = \|A^\top A\|$ , which can be quickly approximated using power iteration, therefore we applied [Algorithm 2](#), with fixed stepsize  $\gamma = 0.95/L_f$ .

In [Figure 4](#) the performance of MinFBE(BFGS) is shown in a small dimensional instance taken from the SPEAR datasets.<sup>2</sup> It is apparent that our method greatly improves over FBS, its accelerated version, and classical BFGS applied to the problem of minimizing  $\varphi_\gamma^{\text{FB}}$ .

Then we considered larger instances from the same dataset. In this case we applied L-BFGS and the nonlinear conjugate gradient method by Dai and Yuan (CG-DY, see [\[42\]](#)), which always produces descent directions when a line-search satisfying the Wolfe conditions is employed. The same formulas were used in the context of MinFBE: in this case CG-DY does not necessarily produce descent directions, therefore we restart the memory of the method every time an ascent direction is encountered. We also compare against SpaRSA [\[161\]](#), a proximal gradient algorithm using the Barzilai-Borwein method to determine the stepsize and a nonmonotone line-search to guarantee convergence, and FPC\_AS [\[157\]](#), which is an active-set type of algorithm. These are *ad-hoc* solvers for  $\ell_1$ -regularization problems, in contrast to our approach which is for general problems of the form [\(2.1\)](#). Both SpaRSA and FPC\_AS adopt a *continuation* strategy to warm-start the problem and accelerate convergence. For the sake of fairness we ran also the other methods (fast FBS, L-BFGS, CG-DY and MinFBE) in a similar continuation scheme: we solve a sequence of problems, with a large initial value of  $\lambda$  (close to  $\lambda_{\max}$ ) which is successively reduced until the target value is reached, using the solution to each problem as initial iterate for the successive. As it is apparent from the results in [Figure 5](#), MinFBE(L-BFGS) and MinFBE(CG-DY) are able to solve all the instances we considered and generally outperform the other methods, including the corresponding classical line-search methods. Therefore, the additional forward-backward step performed by MinFBE after the descent step indeed pays off.

---

<sup>2</sup><http://wwwopt.mathematik.tu-darmstadt.de/spear/>



**Figure 4:** Lasso: algorithms applied to the `spear_inst_1`, with  $m = 512$  samples and  $n = 1024$  variables, where  $\lambda = 0.05\lambda_{\max}$  was used. MinFBE converges superlinearly to the global minimum when BFGS directions are used, and faster than the classical BFGS algorithm applied to the problem of minimizing  $\varphi_{\gamma}^{\text{FB}}$ .

## 2.5.2 Sparse logistic regression

The composite objective function consists of

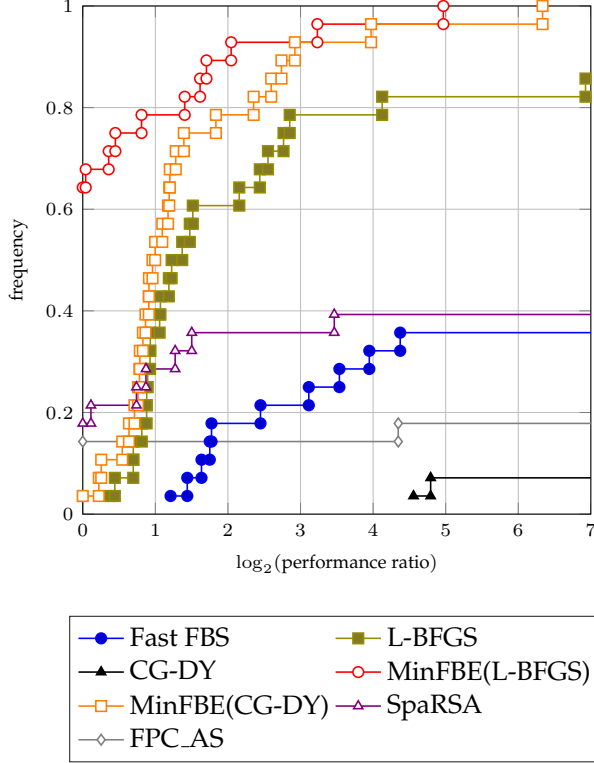
$$f(x) = \sum_{i=1}^m \log(1 + e^{-b_i \langle a_i, x \rangle}), \quad g(x) = \lambda \|x\|_1.$$

Here vector  $a_i \in \mathbb{R}^n$  contains the features of the  $i$ -th instance, and  $b_i \in \{-1, 1\}$  indicates the correspondent class. The  $\ell_1$ -regularization enforces sparsity in the solution. Indicating by  $A$  the matrix having  $a_i$  as  $i$ -th row, we have  $\lambda_{\max} = \frac{1}{2} \|A^\top b\|_\infty$ , so that for  $\lambda \geq \lambda_{\max}$  the optimal solution is  $x_\star = 0$ .

We ran the algorithms on three datasets,<sup>3</sup> and recorded the number of iterations, calls to  $f$  and  $\nabla f$ , matrix-vector products with  $A$  and  $A^\top$ , and the running time needed to reach  $\varphi(x^k) - \varphi_\star \leq 10^{-8}(1 + |\varphi_\star|)$ . Unlike

<sup>3</sup><http://www.csie.ntu.edu.tw/~cjlin/libsvmtools/datasets/>





**Figure 5:** Lasso: performance profile of the CPU time, for the problems in the SPEAR dataset ranging from `spear_inst_173` to `spear_inst_200`, and  $\lambda = 10^{-3}\lambda_{\max}$ . All algorithms use a continuation technique to warm-start the problem solution. Each method was stopped as soon as  $\varphi(x^k) - \varphi_* \leq 10^{-6}(1 + |\varphi_*|)$ . Methods not meeting this condition in  $10^4$  iterations were assigned a performance ratio of  $+\infty$ .

the previous example, here a tight Lipschitz constant for  $\nabla f$  is not readily available: in this case we applied MinFBE (as well as fast FBS) with backtracking on parameter  $\gamma$ . The results are in [Table 2](#): MinFBE significantly reduces the number of operations needed to solve the problems. Since directions are computed according to L-BFGS, which is able to scale to large dimensional problems, CPU time is reduced analogously.

### 2.5.3 Group lasso

Let vector  $x$  be partitioned as  $x = (x_1, \dots, x_N)$ , where each  $x_i \in \mathbb{R}^{n_i}$ , and  $\sum_i n_i = n$ . We consider the  $\ell_2$ -regularized least squares problem having

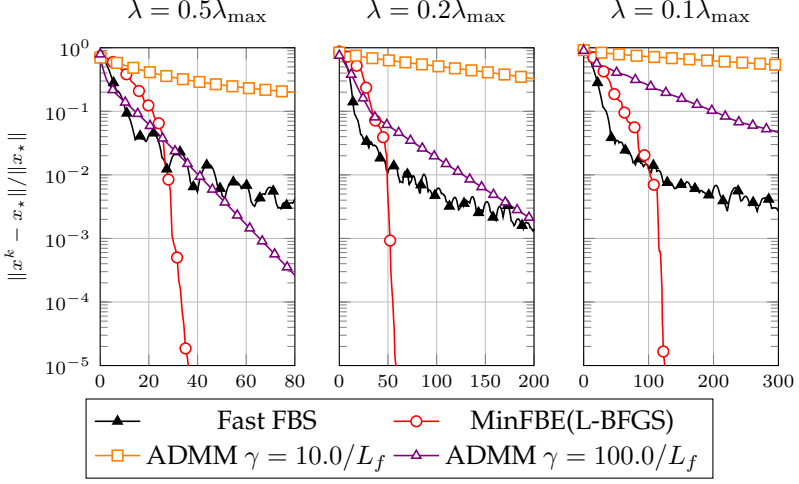
$$f(x) = \frac{1}{2} \|Ax - b\|_2^2, \quad g(x) = \lambda \sum_{i=1}^N \|x_i\|_2.$$

The  $\ell_2$  terms enforce sparsity at the block level, so that for sufficiently large  $\lambda$  we expect many of the  $x_i$ 's to be zero. Partitioning the  $A$  by columns as  $A = (A_1, \dots, A_N)$ , with the same block structure as  $x$ , then for  $\lambda \geq \lambda_{\max} = \max \{\|A_1^\top b\|_2, \dots, \|A_N^\top b\|_2\}$  the optimal solution is  $x_\star = 0$ .

To test the methods we generated a random instance as follows: we set  $m = 200$ ,  $N = 2000$  and  $n_1 = \dots = n_N = 100$ , and generated  $A$  as a sparse matrix with normally distributed entries, density  $10^{-2}$  and condition number  $10^2$  using MATLAB's `sprandn` command. Then we chose  $x_{\text{true}}$  with 10 nonzero blocks, and computed  $b = Ax_{\text{true}} + v$ , where  $v$  is a Gaussian noise vector with standard deviation 0.1. Just like in the case of lasso, the Lipschitz constant  $L_f$  can be easily estimated using power iterations. We compared fast FBS, MinFBE(L-BFGS) and ADMM (with two different stepsize parameters  $\gamma$ ), on such an instance. As it is shown in [Figure 6](#), MinFBE exhibits fast asymptotic convergence, and approaches the solution much faster than fast FBS and ADMM. Unlike ADMM, no tuning of  $\gamma$  is needed in MinFBE to obtain fast convergence.

| ID  | $\lambda/\lambda_{\max}$ | $\text{nnz}(x_*)$ | Fast FBS |       |       | L-BFGS   |      |      | MinFBE(L-BFGS) |          |     |      |      |          |
|---|--------------------------|-------------------|----------|-------|-------|----------|------|------|----------------|----------|-----|------|------|----------|
|   |                          |                   | it.      | $f$   | A     | time (s) | it.  | $f$  | A              | time (s) | it. | $f$  | A    | time (s) |
| <b>rcv1</b><br>$m = 20242$<br>$n = 44504$<br>$\text{nnz}(A) = 910K$     | $2 \cdot 10^{-1}$        | 25                | 134      | 269   | 403   | 1.57     | 58   | 144  | 386            | 1.37     | 29  | 87   | 198  | 0.94     |
|   | $1 \cdot 10^{-1}$        | 70                | 261      | 523   | 784   | 2.91     | 132  | 305  | 843            | 3.68     | 51  | 168  | 367  | 1.46     |
|   | $5 \cdot 10^{-2}$        | 141               | 406      | 813   | 1219  | 4.49     | 170  | 386  | 1075           | 4.65     | 46  | 152  | 332  | 1.30     |
|   | $2 \cdot 10^{-2}$        | 287               | 885      | 1771  | 2656  | 9.75     | 230  | 530  | 1459           | 6.32     | 76  | 239  | 539  | 2.13     |
| <b>real-sim</b><br>$m = 72201$<br>$n = 20958$<br>$\text{nnz}(A) = 1.5M$ | $1 \cdot 10^{-2}$        | 470               | 1189     | 2379  | 3568  | 14.62    | 356  | 787  | 2220           | 8.48     | 105 | 304  | 720  | 2.93     |
|   | $2 \cdot 10^{-1}$        | 19                | 123      | 247   | 370   | 4.62     | 43   | 115  | 296            | 2.86     | 15  | 35   | 91   | 1.38     |
|   | $1 \cdot 10^{-1}$        | 52                | 200      | 401   | 601   | 7.09     | 72   | 176  | 472            | 4.75     | 20  | 56   | 132  | 1.54     |
|   | $5 \cdot 10^{-2}$        | 111               | 325      | 651   | 976   | 14.18    | 93   | 215  | 595            | 5.79     | 29  | 83   | 195  | 2.42     |
| <b>news20</b><br>$m = 19954$<br>$n = 1355191$<br>$\text{nnz}(A) = 3.7M$ | $2 \cdot 10^{-2}$        | 251               | 577      | 1155  | 1732  | 21.05    | 154  | 352  | 976            | 9.46     | 48  | 139  | 327  | 3.42     |
|   | $1 \cdot 10^{-2}$        | 448               | 824      | 1649  | 2473  | 33.46    | 220  | 499  | 1388           | 13.68    | 72  | 227  | 511  | 7.09     |
|   | $2 \cdot 10^{-1}$        | 47                | 793      | 1590  | 2383  | 84.03    | 179  | 427  | 1162           | 50.65    | 79  | 264  | 573  | 32.76    |
|   | $1 \cdot 10^{-1}$        | 98                | 1131     | 2265  | 3396  | 125.86   | 341  | 789  | 2172           | 96.70    | 127 | 401  | 902  | 51.74    |
|   | $5 \cdot 10^{-2}$        | 208               | 3106     | 6216  | 9322  | 320.49   | 409  | 944  | 2599           | 125.49   | 193 | 646  | 1411 | 82.85    |
|   | $2 \cdot 10^{-2}$        | 422               | 6647     | 13298 | 19945 | 673.90   | 1082 | 2481 | 6829           | 352.46   | 440 | 1499 | 3252 | 204.97   |

**Table 2:** Sparse logistic regression: performance of the algorithms on three datasets, for different values of  $\lambda$ . We used  $\varphi(x^{(k)}) - \varphi_* \leq 10^{-8}(1 + |\varphi_*|)$  as termination criterion.



**Figure 6:** Group lasso: performance of the proposed method on a random sparse problem with  $m = 200$  data points and  $n = 2 \cdot 10^5$  variables. Horizontal axis is time in seconds.

### 2.5.4 Matrix completion

We consider the problem of recovering the entries of an  $m$ -by- $n$  matrix, which is known to have small rank, from a sample of them. One may refer to [34] for a detailed theoretical analysis of the problem. The decision variable is now a matrix  $x = (x_{ij}) \in \mathbb{R}^{m \times n}$ , and the problem has the form

$$f(x) = \frac{1}{2} \|\mathcal{A}(x) - b\|^2, \quad g(x) = \lambda \|x\|_*,$$

where  $\mathcal{A} : \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^k$  is a linear mapping selecting  $k$  entries from  $x$ , vector  $b \in \mathbb{R}^k$  contains the known entries, and  $\|x\|_*$  indicates the nuclear norm of  $x$ , which is the sum of its singular values. In this case  $L_f = 1$ , therefore we applied MinFBE with constant  $\gamma = 0.95$ .

The most computationally expensive operation here is the proximal step, requiring a singular value decomposition (SVD). Computing the full SVD becomes infeasible as  $m$  and  $n$  grow, therefore we use the fol-

lowing partial decomposition strategy in evaluating  $\text{prox}_{\gamma g}$ : start with  $\nu_0 = 10$ , and the  $i$ -th time  $\text{prox}_{\gamma g}$  is evaluated compute only the largest  $\nu_i$  singular values  $\sigma_1 \geq \dots \geq \sigma_{\nu_i}$ , and

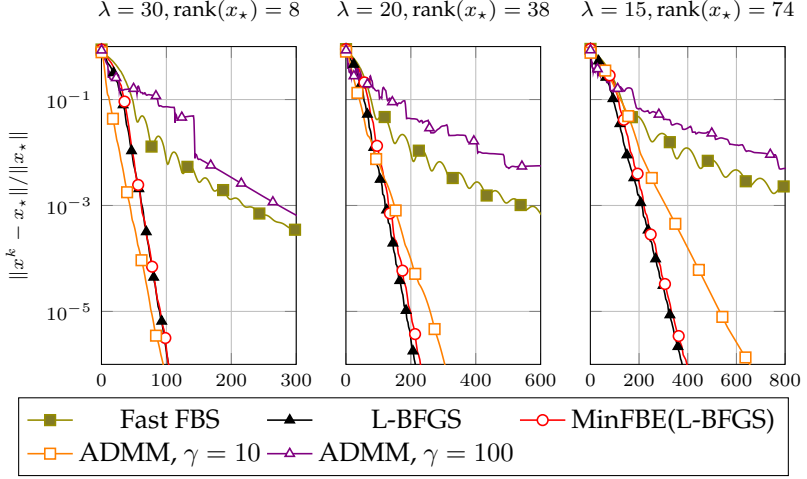
$$\text{prox}_{\gamma g}(x) \approx U \tilde{\Sigma}_+ V^T, \quad \tilde{\Sigma}_+ = \text{diag}(\max\{0, \sigma_i - \gamma\lambda\}, i = 1, \dots, \nu_i).$$

Then set  $\nu_{i+1}$  according to the following rule

$$\nu_{i+1} = \begin{cases} \min\{j \mid \sigma_j \leq \gamma\lambda\} & \text{if } \sigma_{\nu_i} \leq \gamma\lambda \\ \nu_i + 5 & \text{otherwise.} \end{cases}$$

The same technique for approximately thresholding the singular values is used in other algorithms for nuclear norm regularization problems [153]. The partial singular value decompositions were performed using PROPACK software package.<sup>4</sup>

We compared fast FBS, L-BFGS, MinFBE(L-BFGS) and ADMM on the MovieLens100k dataset.<sup>5</sup> This consists of  $10^5$  ratings of 1682 movies from 943 users, so that the problem has  $\approx 1.6$  millions variables. The results of the simulations, for decreasing values of  $\lambda$ , are in Figure 7. Unlike the previous example, in this case MinFBE performs very similarly to standard L-BFGS: they both converge considerably faster than the accelerated FBS, and generally faster than ADMM, especially for smaller values of the regularization parameter. Note also that, just like in the previous example, the performance of ADMM is very sensitive to the value of parameter  $\gamma$ . In our experiment we identified  $\gamma = 10$  as a good value by hand-tuning. Such tuning is not required in MinFBE, where the selection of a suitable  $\gamma$  is automatic.



**Figure 7:** Matrix completion: performance of MinFBE on the MovieLens100k dataset, for different values of  $\lambda$ . Horizontal axis is time in seconds.

## 2.5.5 Image restoration

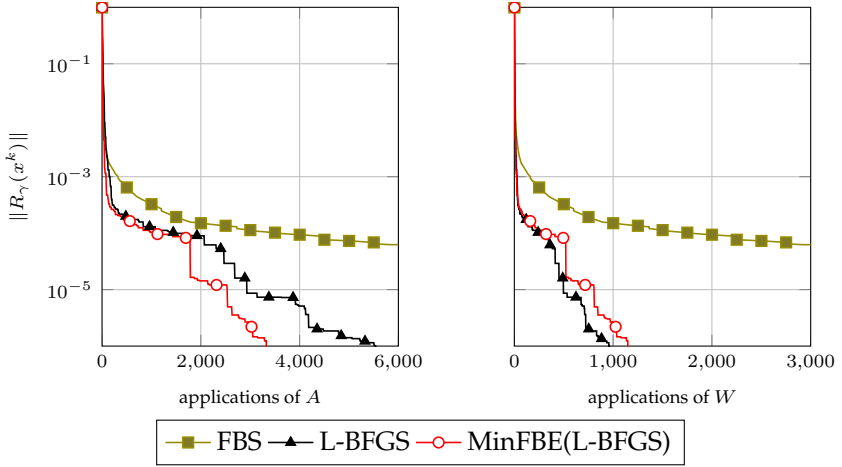
As a nonconvex example we consider the restoration of a noisy blurred  $M \times N$  image. The formulation we use is similar to that in [23], although here we consider the  $\ell_1$  norm in place of the  $\ell_0$  norm as regularization term. Specifically, we set

$$f(x) = \sum_{i=1}^{MN} \psi((Ax - b)_i), \quad g(x) = \lambda \|Wx\|_1.$$

Here,  $b$  denotes the noisy blurred image,  $A$  is a Gaussian blur operator and  $W$  is a discrete Haar wavelet transform with four levels, while  $\psi(t) = \log(1 + t^2)$ , therefore here  $\nabla f$  has Lipschitz constant 2. Since  $W^\top W =$

<sup>4</sup><http://sun.stanford.edu/~rmunk/PROPACK/>

<sup>5</sup><http://grouplens.org/datasets/movielens/>



**Figure 8:** (Nonconvex) image restoration: performance of MinFBE compared with FBS. On the horizontal axis, number of calls to the blur operator (left plot) and Haar operator (right plot); on the vertical axis the fixed-point residual  $R_\gamma$ . Original, noisy/blurred, and recovered images are shown in Figure 9.

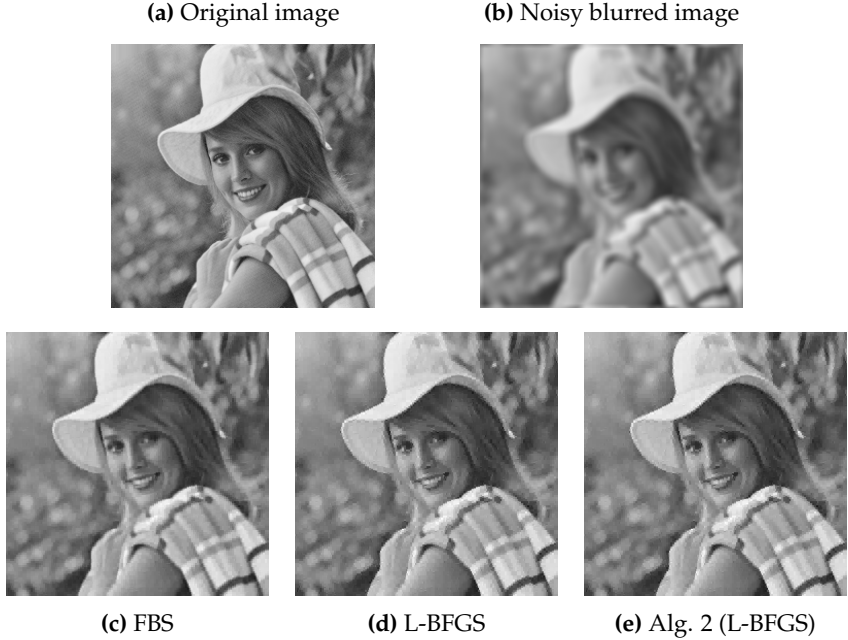
$WW^\top = I$ , the proximal mapping of  $g$  can be computed as

$$\text{prox}_{\gamma g}(x) = W^\top \text{prox}_{\gamma \|\cdot\|_1}(Wx). \quad (2.40)$$

We applied MinFBE to a  $256 \times 256$ -pixel black-and-white image. We distorted the original image with a Gaussian blur operator  $9 \times 9$  with standard deviation 4, and with Gaussian noise with standard deviation  $10^{-3}$ . The regularization parameter in (2.40) was set as  $\lambda = 10^{-4}$ . Results of the simulations are shown in Figures 9 and 8.

## 2.6 Conclusions

The forward-backward splitting (FBS) algorithm for minimizing  $\varphi = f + g$ , where  $f$  is smooth and  $g$  is convex, is equivalent to a variable-metric gradient method applied to a continuously differentiable objec-



**Figure 9:** (Nonconvex) image restoration: recovered images obtained with the three considered algorithms.

tive, which we called forward-backward envelope (FBE), when  $f \in C^2$ . Therefore, we can adopt advanced smooth unconstrained minimization algorithms, such as quasi-Newton and limited-memory methods, to the problem of minimizing the FBE and thus solving the original, nonsmooth problem. We propose to implement them in an algorithmic scheme, which we call MinFBE, which is appealing in that (i) it relies on the very same black-box oracle as FBS (evaluations of  $f$ , its gradient,  $g$  and its proximal mapping) and is therefore suited for large scale applications, (ii) it does not require the knowledge of global information such as Lipschitz constant  $L_f$ , but can adaptively estimate it. The proposed method exploits the composite structure of  $\varphi$ , and alternates line-search steps over descent directions and forward-backward steps. For this reason,



MinFBE possesses the same global convergence properties of FBS, under the assumptions that  $\varphi$  has the Kurdyka-Łojasiewicz properties at its critical points, and a global convergence rate  $O(1/k)$  in case  $\varphi$  is convex. This is a peculiar feature of our approach, since line-search methods do not converge to stationary points, in general, when applied to nonconvex functions. Moreover, we proved that when quasi-Newton directions are used in MinFBE, and the FBE is twice differentiable with nonsingular Hessian at the limit point of the sequence of iterates, superlinear asymptotic convergence is achieved. Our theoretical results are supported by numerical experiments. These show that MinFBE with (limited-memory) quasi-Newton directions improves the asymptotic convergence of FBS (and its accelerated variant when  $\varphi$  is convex), and usually converges faster than the corresponding classical line-search method applied to the problem of minimizing the FBE.

## 2.A Additional results

The next result states that strict differentiability is preserved by composition; its proof is a trivial computation and is therefore omitted.

**Proposition 2.A.1.** *Let  $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ ,  $P : \mathbb{R}^m \rightarrow \mathbb{R}^k$ . Suppose that  $F$  and  $P$  are (strictly) differentiable at  $\bar{x}$  and  $F(\bar{x})$ , respectively. Then the composition  $T = P \circ F$  is (strictly) differentiable at  $\bar{x}$ .*

Using [Proposition 2.A.1](#) one can see that the product of functions (in the matrix-vector sense) preserves (strict) differentiability. However, if one of the two functions vanishes at one point, then we may relax some assumptions, as it is proved in the next result.

**Proposition 2.A.2.** *Let  $Q : \mathbb{R}^n \rightarrow \mathbb{R}^{m \times k}$  and  $R : \mathbb{R}^n \rightarrow \mathbb{R}^k$ , and suppose that  $R(\bar{x}) = 0$ . If  $Q$  is (strictly) continuous at  $\bar{x}$  and  $R$  is (strictly) differentiable at  $\bar{x}$ , then their product  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  defined as  $G(x) = Q(x)R(x)$  is (strictly) differentiable at  $\bar{x}$  with  $JG(\bar{x}) = Q(\bar{x})JR(\bar{x})$ .*

*Proof.* Suppose first that  $Q$  is continuous at  $\bar{x}$  and  $R$  is differentiable at  $\bar{x}$ . Then, expanding  $R(x)$  at  $\bar{x}$  and since  $G(\bar{x}) = 0$ , we obtain

$$\begin{aligned} \frac{G(x) - G(\bar{x}) - Q(\bar{x})JR(\bar{x})[x - \bar{x}]}{\|x - \bar{x}\|} &= \frac{Q(x)R(x) - Q(\bar{x})JR(\bar{x})[x - \bar{x}]}{\|x - \bar{x}\|} \\ &= \frac{(Q(x) - Q(\bar{x}))JR(\bar{x})[x - \bar{x}]}{\|x - \bar{x}\|} \\ &\quad + \frac{o(\|x - \bar{x}\|)}{\|x - \bar{x}\|}. \end{aligned}$$

The quantity  $JR(\bar{x})[\frac{x-\bar{x}}{\|x-\bar{x}\|}]$  is bounded, and continuity of  $Q$  at  $\bar{x}$  implies that taking the limit for  $\bar{x} \neq x \rightarrow \bar{x}$  yields 0. This proves that  $G$  is differentiable at  $\bar{x}$ .

Suppose now that  $Q$  is strictly continuous at  $\bar{x}$ , and that  $R$  is strictly differentiable at  $\bar{x}$ . Then, expanding  $R(y)$  at  $x$  we obtain

$$\begin{aligned} \frac{G(y) - G(x) - Q(\bar{x})JR(\bar{x})[y - x]}{\|y - x\|} &= \frac{(Q(y) - Q(\bar{x}))JR(\bar{x})[y - x]}{\|y - x\|} \\ &\quad + \frac{(Q(y) - Q(x))R(x)}{\|y - x\|} \\ &\quad + \frac{Q(y)o(\|x - y\|)}{\|y - x\|}. \end{aligned}$$

The quantity  $JR(\bar{x})[\frac{y-x}{\|y-x\|}]$  is bounded, and by strict continuity of  $Q$  at  $\bar{x}$  so is  $\frac{Q(x)-Q(y)}{\|x-y\|}$  for  $x, y$  sufficiently close to  $\bar{x}$ . Taking the limit for  $(x, y) \rightarrow (\bar{x}, \bar{x})$  with  $x \neq y$  in the above expression then yields 0, proving strict differentiability. Uniqueness of the Jacobian proves also the claimed form of  $JG(\bar{x})$ .  $\square$

**Proposition 2.A.3.** *Suppose that  $G : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is semidifferentiable in a neighborhood  $N$  of  $\bar{x}$  and that  $DG$  is calm at  $\bar{x}$ , i.e., there exists  $L > 0$  such that, for all  $x \in N$  and  $d \in \mathbb{R}^n$  with  $\|d\| = 1$ ,*

$$\|DG(x)[d] - DG(\bar{x})[d]\| \leq L\|x - \bar{x}\|.$$

Then,

$$\|G(x) - G(y) - DG(\bar{x})[x - y]\| \leq L \max \{\|x - \bar{x}\|, \|y - \bar{x}\|\} \|x - y\|.$$

*Proof.* Follows from [84, Lem. 2.2] by observing that the assumption of Lipschitz-continuity may be relaxed to calmness.  $\square$

The following results are instrumental in proving convergence of the iterates of MinFBE.

**Lemma 2.A.4.** *Under [Assumption 2.1](#), consider the sequences  $(x^k)_{k \in \mathbb{N}}$  and  $(w^k)_{k \in \mathbb{N}}$  generated by [Algorithm 1](#). If there exist  $\bar{\tau}, c > 0$  such that  $\tau_k \leq \bar{\tau}$  and  $\|d^k\| \leq c\|R_{\gamma_k}(x^k)\|$ , then*

$$\|x^{k+1} - x^k\| \leq \gamma_k \|R_{\gamma_k}(w^k)\| + \bar{\tau} c \|R_{\gamma_k}(x^k)\| \quad \forall k \in \mathbb{N} \quad (2.41)$$

and, for  $k$  large enough,

$$\|x^{k+1} - x^k\| \leq \gamma_k \|R_{\gamma_k}(w^k)\| + \bar{\tau} c (1 + \gamma_k L_f) \|R_{\gamma_{k-1}}(w^{k-1})\| \quad (2.42)$$

*Proof.* Equation (2.41) follows simply by

$$\|x^{k+1} - x^k\| = \|x^{k+1} - w^k + \tau_k d^k\| \leq \gamma_k \|R_{\gamma_k}(w^k)\| + \bar{\tau} c \|R_{\gamma_k}(x^k)\|.$$

Now, for  $k$  sufficiently large  $\gamma_k = \gamma_{k-1} = \gamma_\infty > 0$ , see [Lemma 2.3.1](#), and

$$\begin{aligned} \|R_{\gamma_k}(x^k)\| &= \gamma_k^{-1} \|x^k - T_{\gamma_k}(x^k)\| \\ &= \gamma_k^{-1} \|T_{\gamma_k}(w^{k-1}) - T_{\gamma_k}(x^k)\| \\ &\leq \gamma_k^{-1} \|w^{k-1} - \gamma_k \nabla f(w^{k-1}) - x^k + \gamma_k \nabla f(x^k)\| \\ &\leq \gamma_k^{-1} \|w^{k-1} - x^k\| + \|\nabla f(w^{k-1}) - \nabla f(x^k)\| \\ &\leq (1 + \gamma_k L_f) \|R_{\gamma_{k-1}}(w^{k-1})\|, \end{aligned}$$

where the first inequality follows from nonexpansiveness of  $\text{prox}_{\gamma g}$ , and the last one from Lipschitz continuity of  $\nabla f$ . Putting this together with

(2.41) gives (2.42).  $\square$

**Lemma 2.A.5.** *Let  $(\beta_k)_{k \in \mathbb{N}}$  and  $(\delta_k)_{k \in \mathbb{N}}$  be real sequences satisfying  $\beta_k \geq 0$ ,  $\delta_k \geq 0$ ,  $\delta_{k+1} \leq \delta_k$  and  $\beta_{k+1}^2 \leq (\delta_k - \delta_{k+1})\beta_k$  for all  $k \in \mathbb{N}$ . Then  $\sum_{k=0}^{\infty} \beta_k < \infty$ .*

*Proof.* Taking the square root of both sides in  $\beta_{i+1}^2 \leq (\delta_i - \delta_{i+1})\beta_i$  and using

$$\sqrt{\zeta\eta} \leq (\zeta + \eta)/2,$$

for any nonnegative numbers  $\zeta, \eta$ , we arrive at  $2\beta_{i+1} \leq (\delta_i - \delta_{i+1}) + \beta_i$ . Summing up the latter for  $i = 0, \dots, k$ , for any  $k \in \mathbb{N}$ ,

$$\begin{aligned} 2\sum_{i=0}^k \beta_{i+1} &\leq \sum_{i=0}^k (\delta_i - \delta_{i+1}) + \sum_{i=0}^k \beta_i \\ &= \delta_0 - \delta_{k+1} + \beta_0 - \beta_{k+1} + \sum_{i=0}^k \beta_{i+1} \\ &\leq \delta_0 + \beta_0 + \sum_{i=0}^k \beta_{i+1}. \end{aligned}$$

Hence

$$\sum_{i=0}^{\infty} \beta_{i+1} \leq \delta_0 + \beta_0 < \infty, \quad (2.43)$$

which concludes the proof.  $\square$

**Proposition 2.A.6.** *Suppose [Assumption 2.1](#) is satisfied and that  $\varphi$  is lower bounded, and consider the sequences generated by [Algorithm 1](#). If  $\beta \in (0, 1)$  and there exist  $\bar{\tau}, c > 0$  such that  $\tau_k \leq \bar{\tau}$  and  $\|d^k\| \leq c\|R_{\gamma_k}(x^k)\|$  then*

$$\sum_{k=0}^{\infty} \|x^{k+1} - x^k\|^2 < \infty. \quad (2.44)$$

*If moreover  $(x^k)_{k \in \mathbb{N}}$  is bounded, then*

$$\lim_{k \rightarrow \infty} \text{dist}_{\omega(x^0)}(x^k) = 0 \quad (2.45)$$

*and  $\omega(x^0)$  is a nonempty, compact and connected subset of  $\text{zer } \partial\varphi$  over which  $\varphi$  is constant.*

*Proof.* (2.44) follows from (2.41), Propositions 2.3.4(ii) and 2.3.4(iv), and the fact that the sum of square-summable sequences is square summable.

If  $(x^k)_{k \in \mathbb{N}}$  is bounded, that  $\omega(x^0)$  is nonempty, compact and connected and  $\lim_{k \rightarrow \infty} \text{dist}_{\omega(x^0)}(x^k) = 0$  follow by [21, Lem. 5(ii),(iii), Remark 5]. That  $\varphi$  is constant on  $\omega(x^0)$  follows by a similar argument as in [21, Lem. 5(iv)].  $\square$

The following is [21, Lem. 6], therefore we state it with no proof.

**Lemma 2.A.7** (Uniformized KL property). *Let  $K \subset \mathbb{R}^n$  be a compact set and suppose that the proper lower semi-continuous function  $\varphi : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is constant on  $K$  and satisfies the KL property at every  $x^* \in K$ . Then there exist  $\varepsilon > 0$ ,  $\eta > 0$ , and a continuous concave function  $\psi : [0, \eta] \rightarrow [0, +\infty)$  such that properties 2.3.9(i), 2.3.9(ii) and 2.3.9(iii) hold, and*

$$(i') \text{ for all } x_* \in K \text{ and } x \text{ such that } \text{dist}_K(x) < \varepsilon \text{ and } \varphi(x_*) < \varphi(x) < \varphi(x_*) + \eta, \\ \psi'(\varphi(x) - \varphi(x_*)) \text{dist}(0, \partial\varphi(x)) \geq 1. \quad (2.46)$$

## Chapter 3

# A simple and efficient algorithm for nonlinear MPC

### 3.1 Introduction

Model predictive control (MPC) has become a popular strategy to implement feedback control loops for a variety of systems, due to its ability to take into account for constraints on inputs, states and outputs. Its success is intimately tied to the availability of efficient, reliable algorithms for the solution of the underlying constrained optimization problem: linear MPC requires the solution of a convex QP at every sampling step, and for this can count on the mature theory of convex optimization algorithms, that provides simple and robust methods with global convergence guarantees.

On the other hand, the vast majority of systems are nonlinear by nature, and nonlinear models often capture their dynamics much more accurately. For this reason nonlinear MPC (NMPC) is a well suited ap-

proach to design feedback controllers in many cases. At every sampling step, NMPC requires the solution of a general nonlinear program (NLP): general approaches for NLP include sequential quadratic programming (SQP) and interior-point methods (IP) [108, 16]. Typically this NLP represents a discrete-time approximation of the continuous-time, and thus infinite-dimensional, constrained nonlinear optimal control problem, within a direct optimal control framework. Various ways exist for deriving a finite-dimensional NLP from a continuous-time optimal control problem, namely single shooting, multiple shooting and collocation methods, see e.g. [50, 124].

Although multiple shooting formulations (keeping the states as problem variables) are recently quite popular, single shooting formulations (implicitly eliminating the states) have been traditionally been used to exploit the sequential structure in optimal control problems, see [56, 160] and [16, §2.6] for a textbook account.

### 3.1.1 Problems framework and motivation

In this chapter we deal with discrete-time, optimal control problems with nonlinear dynamics. This type of problems can be obtained, for example, by appropriately discretizing continuous-time problems. Furthermore, we allow for nonsmooth (possibly nonconvex) penalties on the inputs: these can be (hard or soft) input constraints, or could be used for example to impose (group) sparsity on the input variables by using sparsity-inducing penalties. Note that problems with soft state constraints fit this framework by including an additional smooth penalty on the system state (e.g., the squared Euclidean distance from a constrained set), in the spirit of a generalized quadratic penalty method.

By eliminating the state variables and expressing the cost as a function of the inputs only (single-shooting formulation), the NMPC problems that we address can be reduced to the minimization of a smooth, nonconvex function  $f$  plus a nonsmooth (possibly nonconvex) penalty  $g$ .

This is precisely the form of problems that can be solved by the proximal gradient method, also known as forward-backward splitting (FBS), see [6], a generalization of the projected gradient method for minimizing a smooth function over a constrained set. FBS method is a fixed-point iteration for solving a nonsmooth, nonlinear system of equations defining the stationary points of the cost function. As such, its iterations are very simple and ideal for embedded applications. However, the simplicity of FBS comes at the cost of slow convergence to stationary points. In fact, as it is the case for all first-order methods, the behaviour of FBS is greatly affected by the problem conditioning: in the case of NMPC, it is customary to have very ill-conditioned problems due to the nonlinear dynamics and the horizon length.

### 3.1.2 Contributions

We propose a new, simple method for solving composite problems. The proposed algorithm is a line-search method for solving the fixed-point equations associated with FBS, using the so-called *forward-backward envelope* (FBE) as merit function to determine the stepsize [113, 145, 150]. The algorithm is simpler than the scheme discussed in [113, 145], in that it does not require computing the gradient of the FBE, therefore no second-order information on  $f$  is required. This is particularly convenient in applications such as NMPC, where evaluating derivatives can become expensive. Nevertheless, we show that if the search directions are computed using quasi-Newton formulas, then the stepsize becomes eventually 1 and the algorithm converges with superlinear asymptotic rate to a stationary point. Computing the directions and the value of the FBE simply requires evaluations of the forward-backward mapping, therefore the proposed algorithm is based on the very same operations as FBS:

1. evaluation of the gradient of the smooth cost, which can be performed using automatic differentiation (AD), such as CasADi [1], in NMPC applications;



2. evaluation of the proximal mapping of the nonsmooth penalty, usually having a very simple closed-form.

Furthermore, limited-memory methods such as L-BFGS [93] that only perform inner products can be used to determine line-search directions, making the algorithm completely matrix-free. These features make the proposed algorithm well suited for embedded implementations and applications.

## 3.2 Newton-type forward-backward method

We target problems of the form

$$\underset{\mathbb{R}^n}{\text{minimize}} \quad \varphi(u) = \ell(u) + g(u), \quad (3.1)$$

where  $\ell$  is smooth and  $g$  is convex. The FBS scheme is based on simple iterations of the form

$$u^{k+1} = T_\gamma(u^k) = \text{prox}_{\gamma g}(u^k - \gamma \nabla \ell(u^k)), \quad (3.2)$$

where  $\gamma > 0$  is a stepsize parameter and  $\text{prox}_{\gamma g}$  is the proximal mapping of  $g$ . First studied for convex problems, FB iterations (3.2) have been recently shown to converge in the nonconvex case [6]: if  $\ell$  is differentiable with  $L_\ell$ -Lipschitz continuous gradient then for any  $\gamma \in (0, 2/L_\ell)$  all cluster points of the sequence generated by (3.2) are critical, meaning that they satisfy the first order necessary condition for optimality. Moreover, if  $\varphi = \ell + g$  has the Kurdyka-Łojasiewicz property – a mild property satisfied by all subanalytic functions, for instance – then any bounded sequence (3.2) is globally convergent to a unique critical point.

Because of all such favorable properties, and the fact that in many problems the proximal mapping is available in closed form, FBS has been employed and studied extensively. However, the downside of such simple algorithm is its slow tail convergence, being it  $Q$ -linear at best and

with  $Q$ -factor typically close to one when the problem is ill-conditioned. Using variable metrics, *e.g.*, coming from Newton-type schemes, can dramatically improve and robustify the convergence, but it comes at the drawback of prohibitively complicating the proximal steps, thus requiring inner procedures possibly as hard as solving the original problem itself. One option in the case of  $g = \delta_C$ , where  $C$  is a box, is to apply the two-metric projection method of Gafni & Bertsekas [69], the trust-region algorithm of [91], or the limited-memory BFGS algorithm for bound constrained optimization in [29], or, more generally, when  $C$  is a simple polyhedral set (one that is easy to project onto) the algorithms of [31]. When  $C$  has a more complicated structure, extensions of this class of methods become quite complex [55]. However, when  $g$  is a nonsmooth function (such as sparsity and group-sparsity inducing penalties) then the aforementioned algorithms are not applicable.

Algorithm 1, presented in Chapter 2, has the unfavourable property of requiring *second-order* information of the smooth term in the cost, which is required to compute the gradient of the FBE in order to obtain descent directions: although this can be approximated using additional gradient evaluations, cf. Remark 2.3.3, this is still an undesirable property especially in the case of NMPC problems (3.1) where gradient evaluations can be computationally very expensive. Relying solely on such a simple black-box oracle as FBS is an extremely appealing property also considering embedded applications and implementations, where the computational power at hand might pose severe limitations to the operations which can be performed.

### 3.2.1 Newton-type methods on generalized equations

Instead of directly addressing the minimization problem, one could target the complementary problem of finding critical points by solving

$$\text{find } u_\star \text{ such that } R_\gamma(u_\star) = 0, \quad (3.3)$$

where

$$R_\gamma(u) = \gamma^{-1} (u - \text{prox}_{\gamma g}(u - \gamma \nabla \ell(u)))$$

is the *fixed-point residual*. This motivates addressing the problem using Newton-type methods

$$u^{k+1} = u^k - H_k R_\gamma(u^k), \quad (3.4)$$

where  $H_k$  are invertible operators that, ideally, capture curvature information of  $R_\gamma$  and enable superlinear or quadratic convergence when close enough to a solution. In quasi-Newton schemes,  $H_k$  is a linear operator recursively updated so as to satisfy the (inverse) secant condition

$$u^{k+1} - u^k = H_{k+1} (R_\gamma(u^{k+1}) - R_\gamma(u^k)),$$

and under mild differentiability assumptions at a candidate limit point  $u_\star$ , local superlinear convergence is achieved provided that the Dennis-Moré condition

$$\lim_{k \rightarrow \infty} \frac{\|R_\gamma(u^k) - JR_\gamma(u_\star)d^k\|}{\|d^k\|} = 0 \quad (3.5)$$

is satisfied, where  $d^k = -H_k R_\gamma(u^k)$ .

### 3.2.2 Forward-backward envelope

The major challenge in designing fast methods for nonlinear equations is making the iterates globally convergent. In fact, without globalization strategies such methods are well known to even possibly diverge. This is usually achieved by *damping* iterations of the form (3.4), premultiplying the update by an appropriate stepsize. In [150] a globalization technique is proposed, based on the *forward-backward envelope* (FBE) [113] (initially derived for convex problems, see also [117, 145]). The FBE is an exact, continuous, real-valued penalty function for the original problem (3.1)

defined as

$$\varphi_\gamma^{\text{FB}}(u) = \ell(u) - \frac{\gamma}{2} \|\nabla \ell(u)\|^2 + g^\gamma(u - \gamma \nabla \ell(u)), \quad (3.6)$$

for a parameter  $\gamma > 0$ . Under no additional assumptions,  $\varphi_\gamma^{\text{FB}}$  enjoys the following favorable properties.

**Proposition 3.2.1.** *Function  $\varphi_\gamma^{\text{FB}}$  is strictly continuous and satisfies*

- (i)  $\varphi_\gamma^{\text{FB}}(u) \leq \varphi(u) - \frac{\gamma}{2} \|R_\gamma(u)\|^2$ ;
- (ii)  $\varphi(T_\gamma(u)) \leq \varphi_\gamma^{\text{FB}}(u) - \frac{\gamma}{2} (1 - \gamma L_\ell) \|R_\gamma(u)\|^2$ .

In particular,

- (iii)  $\varphi(u) = \varphi_\gamma^{\text{FB}}(u)$  for any  $u \in \text{fix } T_\gamma$ ;
- (iv)  $\inf \varphi = \inf \varphi_\gamma^{\text{FB}}$  and  $\text{argmin } \varphi = \text{argmin } \varphi_\gamma^{\text{FB}}$  for any  $\gamma < 1/L_\ell$ .

*Proof.* See [Proposition 2.2.2](#) and [Proposition 2.2.3](#). □

Due to strict continuity, via Rademacher's theorem [[133](#), Theorem 9.60], both  $\nabla \ell$  and  $\varphi_\gamma^{\text{FB}}$  are almost everywhere differentiable with (cf. [Theorem 2.2.6](#))

$$\nabla \varphi_\gamma^{\text{FB}}(u) = Q_\gamma(u) R_\gamma(u), \quad \text{where} \quad Q_\gamma(u) = I - \gamma \nabla^2 \ell(u).$$

Matrix  $Q_\gamma(u)$  is symmetric and defined for almost any  $u$ ; if, additionally,  $\gamma < 1/L_\ell$ , then  $Q_\gamma(u)$  is also positive definite wherever it exists. If  $\ell$  is twice differentiable at a critical point  $u_\star$  and  $\text{prox}_{\gamma g}$  is differentiable at  $u_\star - \gamma \nabla \ell(u_\star)$ , then  $R_\gamma$  is differentiable at  $u_\star$  with Jacobian  $JR_\gamma(u_\star)$ , and  $\varphi_\gamma^{\text{FB}}$  is twice differentiable at  $u_\star$  with Hessian (cf. [Lemma 2.2.9](#) and [Thm. 2.2.10](#))

$$\nabla^2 \varphi_\gamma^{\text{FB}}(u_\star) = Q_\gamma(u_\star) JR_\gamma(u_\star). \quad (3.7)$$

A sufficient condition for  $\text{prox}_{\gamma g}$  to comply with this requirement involves a mild property of twice epi-differentiability of  $g$ , see [[133](#), §13], as required by [Assumption 2.4](#).

---

**Algorithm 3** PANOC (Proximal Averaged Newton-type method for Optimal Control)

---

**Inputs:**  $\gamma \in (0, 1/L_\ell)$ ,  $\sigma \in (0, \frac{\gamma}{2}(1 - \gamma L_\ell))$ ,  $u^0 \in \mathbb{R}^n$

**Initialize:**  $k = 0$

- 1:  $\bar{u}^k = \text{prox}_{\gamma g}(u^k - \gamma \nabla_u \ell(u^k))$ ,  $r^k = \gamma^{-1}(u^k - \bar{u}^k)$
- 2:  $d^k = -H_k r^k$  for some matrix  $H_k \in \mathbb{R}^{n \times n}$
- 3:  $u^{k+1} = u^k - (1 - \tau_k)\gamma r^k + \tau_k d^k$ , where  $\tau_k$  is the largest in  $\{(1/2)^i \mid i \in \mathbb{N}\}$  such that

$$\varphi_\gamma^{\text{FB}}(u^{k+1}) \leq \varphi_\gamma^{\text{FB}}(u^k) - \sigma \|r^k\|^2 \quad (3.8)$$

- 4:  $k \leftarrow k + 1$ , go to step 1
- 

**Theorem 3.2.2.** *Let  $\gamma < 1/L_\ell$  and suppose that  $\nabla \ell$  and  $\text{prox}_{\gamma g}$  are differentiable at a critical point  $u_*$  and at  $u_* - \gamma \nabla \ell(u_*)$ , respectively. Then  $R_\gamma$  is differentiable at  $u_*$  with Jacobian  $JR_\gamma(u_*)$ . Furthermore,  $u_*$  is a strong local minimum for  $\varphi$  iff it is a strong local minimum for  $\varphi_\gamma^{\text{FB}}$ , in which case  $\nabla^2 \varphi_\gamma^{\text{FB}}(u_*)$  is positive definite and  $JR_\gamma(u_*)$  is invertible.*

*Proof.* See [Lem. 2.2.9](#) and [Thm. 2.2.10](#), observing that twice differentiability of  $\ell$  can be assumed to hold only at  $u_*$ .  $\square$

### 3.2.3 A superlinearly convergent algorithm based on FBS

We propose [Algorithm 3](#), which we called PANOC, a new line-search method for problem (3.1) which is even simpler than the one of [150], yet it maintains all the favorable convergence properties. After a quick glance at the favorable properties of the FBE and its kinship with FBS, the methodology of the proposed scheme is elementary. At each iteration, the forward-backward step  $\bar{u}^k$  is computed. Then, a step is taken from  $u^k$  along a convex combination of the “nominal” FBS update direction  $-\gamma r^k$  (which leads to  $\bar{u}^k$ ) and a candidate fast direction  $d^k$ . By doing so, a small enough stepsize  $\tau_k$  exists at every iteration  $k$  that satisfies the sufficient decrease condition (3.8), enabling global convergence. Further-

more when close enough to a solution the fast directions will take over and, under mild assumptions, the scheme will reduce to  $u^{k+1} = u^k + d^k$ . The next results rigorously show these claims.

**Theorem 3.2.3** (Global subsequential convergence). *Consider the iterates generated by PANOC. Then,  $r^k \rightarrow 0$  square-summably, and the sequences  $(u^k)_{k \in \mathbb{N}}$  and  $(\bar{u}^k)_{k \in \mathbb{N}}$  have the same cluster points, all satisfying the necessary condition for local minimality  $u = \text{prox}_{\gamma g}(u - \gamma \nabla \ell(u))$ .*

*Proof.* First, the algorithm is well defined, that is, the line-search (3.8) always terminates in a finite number of backtrackings. In fact, since  $\sigma < \frac{\gamma}{2}(1 - \gamma L_\ell)$  and  $u^{k+1} \rightarrow \bar{u}^k$  as  $\tau_k \searrow 0$ , continuity of  $\varphi_\gamma^{\text{FB}}$ , Prop.s 3.2.1(ii) and 3.2.1(i) imply that for small enough  $\tau_k$  (3.8) holds.

Telescoping (3.8), and since  $\inf \varphi_\gamma^{\text{FB}} = \inf \varphi > -\infty$  we obtain that  $\sum_{k \in \mathbb{N}} \|r^k\|^2 < \infty$ , and in particular that  $r^k \rightarrow 0$ . Suppose now that  $(u^k)_{k \in K} \rightarrow u'$  for some  $u' \in \mathbb{R}^n$  and  $K \subseteq \mathbb{N}$ . Then, since  $\|\bar{u}^k - u^k\| = \gamma \|r^k\| \rightarrow 0$ , in particular  $(\bar{u}^k)_{k \in K} \rightarrow u'$  as well. The converse holds analogously, proving that  $(u^k)_{k \in \mathbb{N}}$  and  $(\bar{u}^k)_{k \in \mathbb{N}}$  have same cluster points. From  $r^k \rightarrow 0$ , since  $r^k = R_\gamma(u^k)$  and the residual mapping  $R_\gamma$  is (Lipschitz) continuous, we also deduce that  $R_\gamma(u') = 0$ , concluding the proof.  $\square$

**Remark 3.2.4** (Lipschitz constant  $L_\ell$ ). In practice, no prior knowledge of the Lipschitz constant  $L_\ell$  is required for PANOC. In fact, replacing  $L_\ell$  with an initial estimate  $L > 0$ , the following instruction can be added right after step 1:

1bis: If  $\ell(\bar{u}^k) > \ell(u^k) - \gamma \langle \nabla \ell(u^k), r^k \rangle + \frac{L}{2} \|\gamma r^k\|^2$

then  $\gamma \leftarrow \gamma/2$ ,  $L \leftarrow 2L$ ,  $\sigma \leftarrow \sigma/2$  and go to step 1.

Whenever the quadratic bound is violated with  $L$  in place of  $L_\ell$ , the estimated Lipschitz constant  $L$  is increased and  $\gamma, \sigma$  are decreased accordingly. Since replacing  $L_\ell$  with any  $L \geq L_\ell$  still satisfies the quadratic Lipschitz upper bound [16, Prop. A.24], it follows that  $L$  is incremented only a finite number of times. Therefore, there exists an iteration  $k_0$  starting from which  $\gamma$  and  $\sigma$  are constant; in particular, all the results of the

chapter remain valid starting from iteration  $k_0$ , at latest.

Moreover, since  $\bar{u}^k \in \text{dom } g$  by construction, if  $g$  has bounded domain and the selected directions  $d^k$  are bounded (as it is the case for any ‘reasonable’ implementation), it suffices that  $\nabla \ell$  is *locally* Lipschitz-continuous (i.e., strictly continuous), and as such any  $\ell \in C^2$  would fit the requirement. In fact, in such case all the sequences  $(u^k)_{k \in \mathbb{N}}$  and  $(\bar{u}^k)_{k \in \mathbb{N}}$  are contained in a compact enlargement  $\Omega$  of  $\text{dom } \partial g$ , and  $L_\ell$  can be then taken as  $\text{lip}_\Omega(\nabla \ell)$ , or adaptively retrieved in practice as indicated above. This is the typical circumstance in (N)MPC where  $g$  encodes constraints on the inputs which in realistic applications cannot be unbounded.  $\square$

As to the computational cost, each evaluation of  $\varphi_\gamma^{\text{FB}}$  in the left-hand side of the line-search condition (3.8) requires exactly the computation of one forward-backward step;  $\varphi_\gamma^{\text{FB}}(u^k)$  on the right-hand side, instead, is available from the previous iteration. In particular, in the best case of stepsize  $\tau_k = 1$  being accepted, each iteration requires exactly one forward-backward step. Under mild favorable assumptions, this is the case in which the directions satisfy the Dennis-Moré criterion for super-linear convergence (3.5), as shown in [Theorem 3.2.5](#) later on.

The Dennis-Moré condition is enjoyed by directions generated with quasi-Newton schemes under differentiability assumptions at the limit point. The rank-two updates of BFGS are the most widely used in smooth optimization, and thanks to the globalization guarantees of PANOC they straightforwardly fit to our framework. Because of problem size, the limited-memory variant L-BFGS is mostly used, which does not require the computation or storage of full matrices  $H_k$  but simply keeps memory of a small number of last pairs  $s_k = u^{k+1} - u^k$  and  $y^k = r^{k+1} - r^k$ , and retrieves  $d = -H_k r^k$  by simply performing scalar products. In [Section 3.4](#) we will show the efficiency of PANOC with L-BFGS directions compared to plain FBS and state-of-the-art solvers.

The next theorem shows that the FBE does not prevent superlinear convergence to take place whenever Newton-type directions are used in

PANOC; namely unit stepsize is eventually accepted and PANOC eventually produces iterates coming through (3.4). This is in stark contrast with the well-known drawback of classical nonsmooth exact penalty functions (the so-called Maratos effect, see, e.g., [16, §5.3]).

**Theorem 3.2.5** (Superlinear convergence). *Suppose that the iterates generated by PANOC converge to a strong local minimum  $u_\star$  of  $\varphi$  at which  $R_\gamma$  and  $\nabla\varphi_\gamma^{\text{FB}}$  are strictly differentiable. If the operators  $(H_k)_{k \in \mathbb{N}}$  satisfy the Dennis-Moré condition (3.5), then eventually stepsize 1 is always accepted and  $u^k \rightarrow u_\star$  at superlinear rate.*

*Proof.* From Thm. 3.2.2 we know that  $G_\star = \nabla^2\varphi_\gamma^{\text{FB}}(u_\star) \succ 0$  and that  $JR_\gamma(u_\star)$  is nonsingular. Since  $\bar{u}^k$  and  $u^k$  converge to  $u_\star$ , up to an index shifting we may assume that  $(u^k)_{k \in \mathbb{N}}$  is contained in an open set in which  $\varphi_\gamma^{\text{FB}}$  is differentiable and  $R_\gamma$  continuous. Since  $r^k = R_\gamma(u^k) \rightarrow 0$ , from (3.5) it follows that  $d^k \rightarrow 0$ . Let  $u_0^{k+1} = u^k + d^k$ . By adding and subtracting  $R_\gamma(u_0^{k+1})$  in the numerator of (3.5), by strict differentiability of  $R_\gamma$  at  $u_\star$  we obtain

$$\lim_{k \rightarrow \infty} \frac{\|R_\gamma(u_0^{k+1})\|}{\|d^k\|} = \lim_{k \rightarrow \infty} \frac{\|R_\gamma(u_0^{k+1})\|}{\|u_0^{k+1} - u^k\|} = 0. \quad (3.9)$$

Since  $JR_\gamma(x^\star)$  is nonsingular and  $u_0^{k+1} \rightarrow u_\star$ , there exists a constant  $\alpha > 0$  such that  $\|R_\gamma(u_0^{k+1})\| \geq \alpha\|u_0^{k+1} - u_\star\|$  for  $k$  large enough. Combined with (3.9) we obtain

$$0 \leftarrow \frac{\|u_0^{k+1} - u_\star\|}{\|u_0^{k+1} - u^k\|} \geq \frac{\|u_0^{k+1} - u_\star\|}{\|u_0^{k+1} - u_\star\| + \|u^k - u_\star\|}.$$

Dividing numerator and denominator by  $\|u^k - u_\star\|$  yields

$$\lim_{k \rightarrow \infty} \frac{\|u^k + d^k - u_\star\|}{\|u^k - u_\star\|} = \lim_{k \rightarrow \infty} \frac{\|u_0^{k+1} - u_\star\|}{\|u^k - u_\star\|} = 0. \quad (3.10)$$



Therefore,

$$\begin{aligned}
\varepsilon_k &= \frac{\varphi_\gamma^{\text{FB}}(u_0^{k+1}) - \varphi_\gamma^{\text{FB}}(u_\star)}{\varphi_\gamma^{\text{FB}}(u^k) - \varphi_\gamma^{\text{FB}}(u_\star)} \\
&= \frac{\frac{1}{2} \langle G_\star(u_0^{k+1} - u_\star), u_0^{k+1} - u_\star \rangle + o(\|u_0^{k+1} - u_\star\|^2)}{\frac{1}{2} \langle G_\star(u^k - u_\star), u^k - u_\star \rangle + o(\|u^k - u_\star\|^2)} \\
&\leq \frac{\|G_\star\| \left( \frac{\|u_0^{k+1} - u_\star\|}{\|u^k - u_\star\|} \right)^2 + \left( \frac{o(\|u_0^{k+1} - u_\star\|)}{\|u^k - u_\star\|} \right)^2}{\lambda_{\min}(G_\star) + \left( \frac{o(\|u^k - u_\star\|)}{\|u^k - u_\star\|} \right)^2} \rightarrow 0
\end{aligned}$$

as  $k \rightarrow \infty$ . Moreover, since  $\bar{u}^k \rightarrow u_\star$  and  $u_\star$  is a (strong) local minimum, eventually  $\varphi_\gamma^{\text{FB}}(\bar{u}^k) \geq \varphi_\gamma^{\text{FB}}(u_\star)$ ; combining with [Prop. 3.2.1\(i\)](#) we obtain

$$\varphi_\gamma^{\text{FB}}(u^k) - \varphi_\gamma^{\text{FB}}(u_\star) \geq \varphi_\gamma^{\text{FB}}(u^k) - \varphi_\gamma^{\text{FB}}(\bar{u}^k) \geq \frac{\gamma}{2}(1 - \gamma L_f) \|r^k\|^2.$$

Therefore,

$$\begin{aligned}
\varphi_\gamma^{\text{FB}}(u_0^{k+1}) - \varphi_\gamma^{\text{FB}}(u^k) &\leq -(1 - \varepsilon_k)(\varphi_\gamma^{\text{FB}}(u^k) - \varphi_\gamma^{\text{FB}}(u_\star)) \\
&\leq -(1 - \varepsilon_k) \frac{\gamma}{2} (1 - \gamma L_f) \|r^k\|^2 \\
&\leq -\sigma \|r^k\|^2 \quad \text{for } k \text{ large enough,}
\end{aligned}$$

where the last inequality follows from the fact that  $\varepsilon_k \rightarrow 0$  and  $\sigma < \frac{\gamma}{2}(1 - \gamma L_f)$ , so that eventually  $(1 - \varepsilon_k) \frac{\gamma}{2} (1 - \gamma L_f) \geq \sigma$ . Therefore, for large enough  $k$  the line-search condition [\(3.8\)](#) holds with  $\tau_k = 1$ , and unitary step-size is always accepted. In particular, the limit [\(3.10\)](#) reads  $\lim_{k \rightarrow \infty} \frac{\|u^{k+1} - u_\star\|}{\|u^k - u_\star\|} = 0$ , proving  $(u^k)_{k \in \mathbb{N}}$  to be superlinearly convergent.  $\square$

It is important to remark again that (strict) differentiability of  $R_\gamma$  and  $\nabla \varphi_\gamma^{\text{FB}}$  at critical points does not require any smoothness condition on the nonsmooth function  $g$ : in fact, by [Lemma 2.2.9](#) and [Thm. 2.2.10](#), it is sufficient that  $g$  has generalized quadratic *strict second epi-derivative* at the limit point, for the assumptions of [Theorem 3.2.5](#) to hold. As discussed in [\[130, 131, 118, 119\]](#), examples where this happens are abundant and

include several functions used in practice.

### 3.3 Nonlinear MPC

We consider the following finite-horizon optimal control problem

$$\text{minimize } \sum_{j=0}^{N-1} \ell_j(x_j, u_j) + g_j(u_j) + \ell_N(x_N) \quad (3.11a)$$

$$\text{subject to } x_0 = \bar{x} \quad (3.11b)$$

$$x_{j+1} = f_j(x_j, u_j), \quad j = 0, \dots, N-1 \quad (3.11c)$$

where  $f_j : \mathbb{R}^{n_x \times n_u} \rightarrow \mathbb{R}^{n_x}$ ,  $j = 0, \dots, N-1$  are smooth mappings representing system dynamics  $\ell_j : \mathbb{R}^{n_x \times n_u} \rightarrow \mathbb{R}$ ,  $j = 0, \dots, N-1$ , and  $\ell_N : \mathbb{R}^{n_x} \rightarrow \mathbb{R}$  are smooth functions representing stage and terminal costs respectively, and  $g_j : \mathbb{R}^{n_u} \rightarrow \mathbb{R}$ ,  $j = 0, \dots, N-1$ , are possibly nonsmooth and extended real-valued functions representing penalties on the inputs, *e.g.*, constraints.

We are interested in simple algorithms for solving (3.11), *i.e.*, algorithms that do not involve a doubly iterative procedure, such as SQP-like methods. One such algorithm is certainly forward-backward splitting (FBS), also known as the proximal gradient method. In particular, let the mapping  $F : \mathbb{R}^{Nn_u} \rightarrow \mathbb{R}^{(N+1)n_x}$  be defined as

$$F(u_0, \dots, u_{N-1}) = (F_0(u_0), \dots, F_N(u)),$$

where

$$\begin{aligned} F_0(u) &= \bar{x} \\ F_{j+1}(u) &= f_j(F_j(u), u_j), \quad j = 0, \dots, N-1, \end{aligned}$$

---

**Algorithm 4** Backward AD

---

**Inputs:**  $x_0 \in \mathbb{R}^n$ ,  $u = (u_0, \dots, u_{N-1})$   
**Initialize:**  $\ell(u) = 0$   
1: **for**  $j = 0, \dots, N - 1$  **do**  
2:    $x_{j+1} = f_j(x_j, u_j)$   
3:    $\ell(u) \leftarrow \ell(u) + \ell_j(x_j, u_j)$   
4: **end for**  
5:  $\ell(u) \leftarrow \ell(u) + \ell_N(x_N)$   
6:  $p_N = \nabla \ell_N(x_N)$   
7: **for**  $j = N - 1, \dots, 0$  **do**  
8:    $p_j = \nabla_{x_j} f_j(x_j, u_j) p_{j+1} + \nabla_{x_j} \ell_j(x_j, u_j)$   
9:    $\nabla_{u_j} \ell(u) = \nabla_{u_j} f_j(x_j, u_j) p_{j+1} + \nabla_{u_j} \ell_j(x_j, u_j)$   
10: **end for**

---

and, denoting  $u = (u_0, \dots, u_{N-1})$ ,

$$\begin{aligned}\ell(u) &= \sum_{j=0}^{N-1} \ell_j(F_j(u), u_j) + \ell_N(F_N(u)), \\ g(u) &= \sum_{j=0}^{N-1} g_j(u_j).\end{aligned}$$

Then, problem (3.11) has the form (3.1) with  $n = Nn_u$ .

The gradient of  $\ell$  can be efficiently calculated by backward automatic differentiation (AD) (also known as reverse differentiation, reverse mode AD, adjoint method, or backpropagation), as in [Algorithm 4](#) [56].

### 3.3.1 Handling state constraints

Many optimal control problems in practice include constraints on the states. The situation is covered by the following more general optimal

control formulation:

$$\begin{aligned}
& \text{minimize} \quad \sum_{j=0}^{N-1} \left[ \ell_j(x_j, u_j) + g_j(u_j) + h_j(C_j(x_j, u_j)) \right] \\
& \quad + \ell_N(x_N) + h_N(C_N(x_N)) \\
& \text{subject to } x_0 = \bar{x} \\
& \quad x_{j+1} = f_j(x_j, u_j), \quad j = 0, \dots, N-1
\end{aligned}$$

where, additionally to problem (3.11),  $h_j : \mathbb{R}^{m_j} \rightarrow \overline{\mathbb{R}}$ ,  $n = 0, \dots, N$  are proper, closed, convex functions with easily computable proximal mapping and  $C_j : \mathbb{R}^{n_x} \times \mathbb{R}^{n_u} \rightarrow \mathbb{R}^{m_j}$ ,  $n \in 0, \dots, N-1$ , and  $C_N : \mathbb{R}^{n_x} \rightarrow \mathbb{R}^{m_N}$  are smooth mappings. For example, when  $h_j$  are indicators of the non-positive orthant then we are left with a classical state-constrained optimal control problem.

Next, consider  $G : \mathbb{R}^{Nn_u} \rightarrow \mathbb{R}^{m_0} \times \dots \times \mathbb{R}^{m_N}$  defined as

$$G(u_0, \dots, u_{N-1}) = (G_0(u), \dots, G_N(u)),$$

where  $G_0 = x_0$ ,

$$G_j(u) = C_j(F_j(u), u_j), \quad j = 1, \dots, N-1,$$

$G_N(u) = C_N(F_N(u))$ , and  $h : \mathbb{R}^{m_0} \times \dots \times \mathbb{R}^{m_N} \rightarrow \overline{\mathbb{R}}$  with

$$h(z) = \sum_{j=0}^{N-1} h_j(z_j) + h_N(z_N).$$

The problem can now be expressed as

$$\text{minimize } \ell(u) + g(u) + h(G(u)).$$

A standard practice in MPC is to include state constraints in the cost function via penalties. The reason for doing so is to avoid the unpleasant

situation of ending up with an infeasible optimal control problem which can easily happen in practice due to disturbances and plant-model mismatch. The usual way of doing so is by relaxing state constraints using a quadratic penalty. Taking this approach one step further, we smoothen out  $h$  by replacing it with its *Moreau envelope*  $h^{1/\mu}$ , i.e., the value function of the parametric problem involved in the definition of the proximal mapping. Here  $\mu$  acts as a penalty parameter: in the case of state constraints of the form  $G(u) \in C$ , one has  $h^{1/\mu}(G(u)) = \frac{\mu}{2} \text{dist}_C^2(G(u))$  and the larger the value of  $\mu$ , the larger is the penalty one has to pay for violating the state constraints.

It is well known that the Moreau envelope is smooth whenever  $h$  is a proper, closed, convex function. In fact its gradient is given by

$$\nabla h^{1/\mu}(z) = \mu(z - \text{prox}_{h/\mu}(z)).$$

Since  $G$  is also smooth (as the composition of smooth mappings), the modified stage costs

$$\begin{aligned}\tilde{\ell}_j(u) &= \ell_j(F_j(u), u_j) + h_j^{1/\mu}(G_j(u_j)), \quad j = 0, \dots, N-1 \\ \tilde{\ell}_N(u) &= \ell_N(F_N(u)) + h_N^{1/\mu}(G_N(u_N))\end{aligned}$$

are also smooth, and the same holds for the total cost, which we redefine as

$$\ell \leftarrow \sum_{j=0}^N \tilde{\ell}_j.$$

Therefore, the soft-state-constrained problem has exactly the same form as problem (3.1). The gradient of  $\ell$  can be efficiently computed by extending [Algorithm 4](#) as follows:

**Remark 3.3.1.** For simplicity, in the above discussion we have considered the case where parameter  $\mu$  is a scalar: in case functions  $h_j$  are separable, it is straightforward to adapt [Algorithm 5](#) to the case where  $\mu$  is a vector of parameters, of dimension compatible with the separability structure

---

**Algorithm 5** Backward AD for the cost with soft-constrained states

---

**Inputs:**  $x_0 \in \mathbb{R}^n$ ,  $u = (u_0, \dots, u_{N-1})$   
**Initialize:**  $\ell(u) = 0$

- 1: **for**  $j = 0, \dots, N - 1$  **do**
- 2:    $s_j = \text{prox}_{h_j/\mu}(C_j(x_j, u_j))$
- 3:    $q_j = \mu(C_j(x_j, u_j) - s_j)$
- 4:    $\ell(u) \leftarrow \ell(u) + \ell_j(x_j, u_j) + h(s_j) + \frac{1}{2\mu}\|q_j\|^2$
- 5:    $x_{j+1} = f_j(x_j, u_j)$
- 6: **end for**
- 7:  $s_N = \text{prox}_{h_N/\mu}(C_N(x_N))$
- 8:  $q_N = \mu(C_N(x_N) - s_N)$
- 9:  $\ell(u) \leftarrow \ell(u) + \ell_N(x_N) + h_N(s_N) + \frac{1}{2\mu}\|q_N\|^2$
- 10:  $p_N = \nabla \ell_N(x_N) + \nabla C(x_N)q_N$
- 11: **for**  $j = N - 1, \dots, 0$  **do**
- 12:    $p_j = \nabla_{x_j} f_j(x_j, u_j)p_{j+1} + \nabla_{x_j} \ell_j(x_j, u_j) + \nabla_{x_j} C_j(x_j, u_j)q_j$
- 13:    $\nabla_{u_j} \ell(u) = \nabla_{u_j} f_j(x_j, u_j)p_{j+1} + \nabla_{u_j} \ell_j(x_j, u_j) + \nabla_{u_j} C_j(x_j, u_j)q_j$
- 14: **end for**

---

of  $h_j$ . Similarly, parameter  $\mu$  can be allowed to vary along the stage costs and depend on  $j$ .  $\square$

## 3.4 Numerical Simulations

To test the efficacy of the proposed algorithm we consider a system composed of a sequence of masses connected by springs. The system is analogously described in [158, 156].

The chain is composed by  $M$  masses: one end is connected to the origin, while a handle on the other end allows to control the chain. Let us denote by  $x^i(t) \in \mathbb{R}^3$  the position of the  $i$ -th mass at time  $t$ , for  $i = 1, \dots, M + 1$ , where  $x^{M+1}(t)$  is the position of the control handle. The control action at each time instant is denoted as  $u(t) = \dot{x}^{M+1}(t) \in \mathbb{R}^3$ , i.e., we control the velocity of the handle. Each body in the chain has mass  $m$ , and the springs have constant  $D$  and rest length  $L$ . By Hook's law we

obtain the following dynamics [158]:

$$\begin{aligned}\ddot{x}^i &= \frac{1}{m}(F_{i,i+1} - F_{i-1,i}) + a, \\ F_{i,i+1} &= D \left(1 - \frac{L}{\|x^{i+1} - x^i\|}\right) (x^{i+1} - x^i).\end{aligned}$$

where  $a = (0, 0, -9.81)$  is the acceleration due to gravity. Denoting  $v^i$  the velocity of mass  $i$ , the state vector is

$$x(t) = (x^1(t), \dots, x^{M+1}(t), v^1(t), \dots, v^M(t)).$$

Since we operate in the three-dimensional space, the system has then  $n_x = 3(2M + 1)$  states and  $n_u = 3$  inputs, and its dynamics obeys

$$\dot{x} = f_c(x, u) = (v^1, \dots, v^M, u, \ddot{x}^1, \dots, \ddot{x}^M).$$

### 3.4.1 Simulation scenario

An equilibrium state of the system was computed with the control handle positioned at a given  $x_{\text{end}} \in \mathbb{R}^3$ . This was perturbed by applying a constant input  $u = (-1, 1, 1)$  for 1 second, to obtain the starting position of the chain. The goal is to drive the system back to the reference equilibrium state: this can be achieved by solving, for  $T > 0$

$$\begin{aligned}\underset{x, u}{\text{minimize}} \quad & L_c(T) = \int_0^T \ell_c(x(t), u(t)) dt \\ \text{subject to} \quad & \dot{x} = f_c(x, u)\end{aligned} \tag{3.12}$$

where

$$\ell_c(x, u) = \beta \|x^{M+1} - x_{\text{end}}\|_2^2 + \gamma \sum_{i=1}^M \|v^i\|_2^2 + \delta \|u\|_2^2. \tag{3.13}$$

To discretize (3.12) we consider a sampling time  $t_s$  such that  $T = Nt_s$  and assume that the input  $u$  is piecewise constant according to such

sampling: for  $j = 0, \dots, N - 1$ ,

$$u(t) = u_j, \quad \forall t \in [jt_s, (j+1)t_s).$$

Then

$$L_c(T) = \sum_{j=0}^{N-1} \int_{jt_s}^{(j+1)t_s} \ell_c(x(t), u_j) dt.$$

The problem is cast into the form (3.11) by approximating the above integrals, and the system dynamics, with an appropriate discretization formula and setting

$$\ell_j(x_j, u_j) \approx \int_{nt_s}^{(j+1)t_s} \ell_c(x(t), u_j) dt, \quad (3.14a)$$

$$f_j(x_j, u_j) \approx \int_{nt_s}^{(j+1)t_s} f_c(x(t), u_j) dt, \quad (3.14b)$$

with the initial condition  $x(jt_s) = x_j$ ,  $j = 0, \dots, N - 1$ . Furthermore, we constrain the states and inputs by setting  $g_j$  and  $h_j$  as the indicator functions of the feasible sets as

$$g_j(u) = \delta_{\|\cdot\|_\infty \leq 1}(u),$$

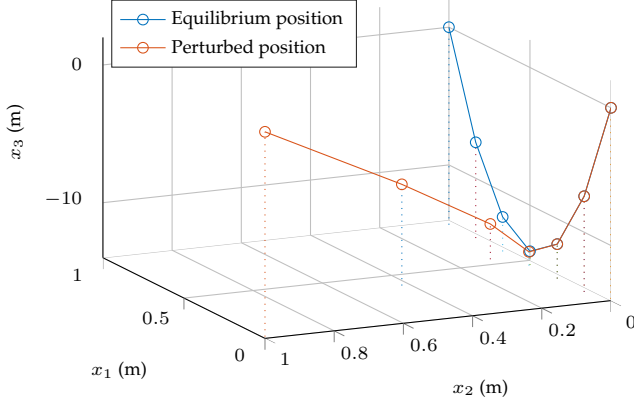
$$h_j(C_j(x, u)) = \sum_{i=1}^{M+1} \delta_{\geq -0.1}(x_2^i).$$

Since  $h_j$  is separable with respect to the different masses, we smoothen it by associating a parameter  $\mu_i$  to each component (see [Section 3.3.1](#) and [Remark 3.3.1](#) in particular):

$$h_j^{1/\mu}(C_j(x, u)) = \sum_{i=1}^{M+1} \frac{\mu_i}{2} (\min \{0, x_2^i + 0.1\})^2. \quad (3.15)$$

In the simulations we have used  $T = 4$  seconds and a sampling time  $t_s = 0.1$  seconds, which gives a prediction horizon  $N = 40$ . The inte-





**Figure 10:** The initial position for the simulation, in the spring-mass system, was obtained perturbing a reference equilibrium state by applying a constant input for 1 second.

grals (3.14) were approximated with a one-step 4th-order Runge-Kutta method. We have used CasADi [1] to implement the discretized dynamics and cost function, so as to efficiently evaluate their Jacobian and gradient. The model parameters were set as  $M = 5$ ,  $m = 0.03$  (kg),  $D = 0.1$  (N/m), and  $L = 0.033$  (m). In the cost function (3.13) we set  $\beta = 1$ ,  $\gamma = 1$ , and  $\delta = 0.01$ . The coefficients for the soft state constraints (3.15) were set as  $\mu_1 = \mu_2 = \mu_3 = 10^2$ ,  $\mu_4 = \mu_5 = \mu_6 = 10$ .

### 3.4.2 Results

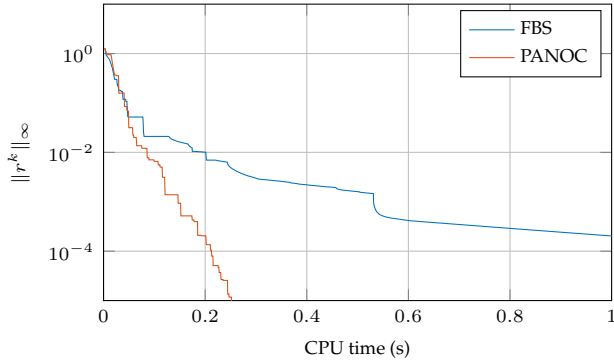
We simulated the system for 15 seconds, with sampling time  $t_s$ , using different solvers to compute the problem solution at every sampling step. To solve the problem we considered PANOC, where the direction  $d^k = -H_k r^k$  in step 2 is computed using the L-BFGS method with memory 10 (see the discussion in Section 3.2.3). Furthermore, we applied FBS, MATLAB's FMINCON (using an SQP algorithm), IPOPT (interior-point method) to the single-shooting formulation. We also applied IPOPT to the multiple-shooting formulation, that gave similar performance, and

to the problem with hard state constraints. We did not apply FMINCON to the multiple-shooting problem, as doing so performed considerably worse. In Figure 11 the convergence of the fixed-point residual  $\|r^k\|_\infty$  for FBS and PANOC is shown for the first problem of the sequence. There we have solved the problem to medium/high accuracy for comparison purposes: in practice, we have noticed that good closed loop performance is obtained with more moderate accuracy. Therefore we ran closed-loop simulations terminating PANOC and FBS as soon as  $\|r^k\|_\infty \leq 10^{-3}$ , while the other solvers were run with default options. The CPU times during the simulation are shown in Figure 12. The effect on the dynamics of soft state constraints is shown in Figure 13, where the trajectory of two masses during the simulation is compared to the hard-constrained and unconstrained cases, the latter obtained by setting  $\mu_i = 0$  for  $i = 1, \dots, M + 1$  in (3.15). It is apparent that by using the soft state constrained formulation, we are able to improve considerably the solution time of the problem, without sacrificing closed loop performance. The proposed PANOC algorithm outperforms all the other considered methods in this example, and greatly accelerates over FBS. This fact is particularly evident in the early stages of the simulation, when the system is far from equilibrium.

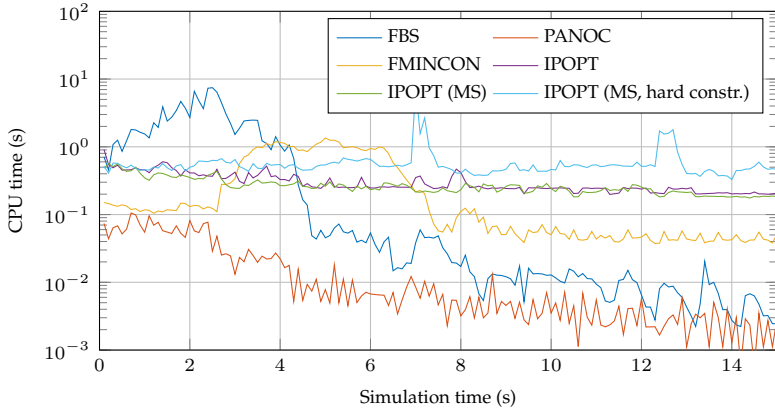
### 3.5 Conclusions

This chapter presents PANOC, a new algorithm for solving nonlinear constrained optimal control problems typically arising in MPC. The algorithm is simple, exploits problem structure, does not require solution of a quadratic program at every iteration and yet can be shown to be superlinearly convergent under mild assumptions. A version of the algorithm that uses simple L-BFGS directions was shown to perform favorably against state-of-the-art NLP solvers in a benchmark example.

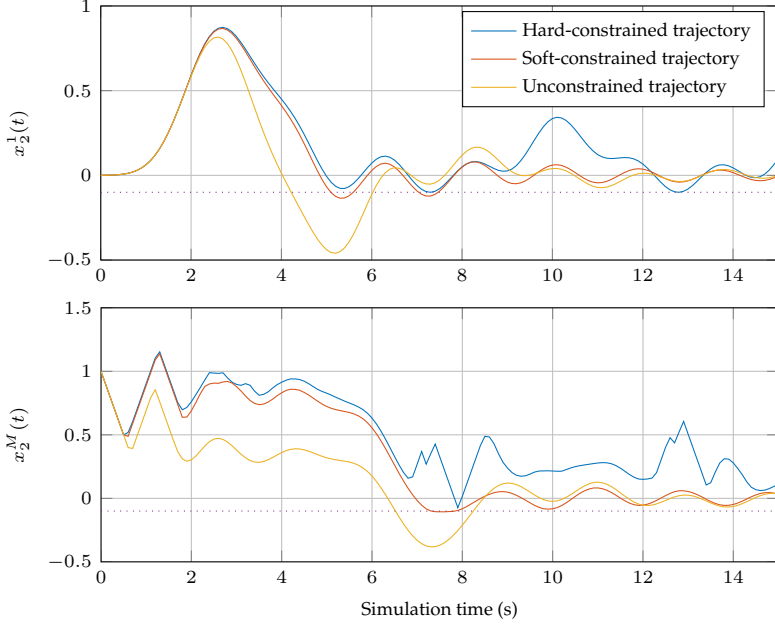
There are several topics for future research: (i) semismooth Newton



**Figure 11:** Convergence of the fixed-point residual for FBS and PANOC, in the first problem of the closed-loop simulation. The algorithms were executed here to medium/high accuracy for comparison purposes.



**Figure 12:** CPU times of different solvers applied to the spring-mass system, during 15 seconds of simulation. “MS” indicates that a solver was applied to the multiple-shooting formulation, where the system state is kept as decision variable. IPOPT was also applied to the problem with hard state constraints for reference.



**Figure 13:** Effect of the soft state constraint terms on the trajectory of the masses 1 and  $M$  in the closed loop simulation. The unconstrained trajectory is obtained by setting  $\mu_i = 0$  for  $i = 1, \dots, M + 1$  in (3.15).

directions [117] that fully exploit problem structure and enable quadratic convergence rates, (ii) more rigorous handling of state constraints by embedding the algorithm in a proximal augmented Lagrangian framework, (iii) a real-time iteration scheme where the algorithm is warm-started by exploiting sensitivity information for the fixed point residual and (iv) development of a code generation tool for embedded applications.

## Chapter 4

# Newton-type alternating minimization algorithm

### 4.1 Introduction

We consider convex optimization problems of the form

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + g(Ax), \tag{4.P}$$

where  $f$  is strongly convex,  $g$  is convex and  $A$  is a linear mapping. Problems of this form are quite general and appear in various areas of applications, including optimal control [144], system identification [60] and machine learning [24, 112]. For example, whenever  $g = \delta_C$  is the indicator function of a convex set  $C$ , then (4.P) models a constrained convex problem: if  $C$  is the nonnegative orthant, then in particular (4.P) amounts to minimizing a strongly convex function subject to polyhedral constraints.

A general approach to the solution of (4.P) is based on the dual proximal gradient method, or forward-backward splitting, also known as alternating minimization algorithm (AMA) [154]. This is the dual appli-

cation of an algorithm introduced by Gabay [68] for finding the zero of the sum of two maximal monotone operators, one of which is assumed to be strongly monotone. The alternating minimization algorithm is intimately tied to the framework of *augmented Lagrangian* methods, and its global convergence and complexity bounds are well covered in the literature, see [154]: a global convergence rate of order  $O(1/\sqrt{k})$  holds for the primal iterates of AMA under very general assumptions, and can be improved to the optimal rate  $O(1/k)$  using a simple acceleration technique due to Nesterov, see [10, 106, 11].

As with all first order methods, the performance of (fast) AMA is severely affected by ill-conditioning of the problem [144]. One way to deal with this issue, which is extensively used in classical smooth, unconstrained optimization, is to precondition the problem using (approximate) second order information on the cost function, as in (quasi-) Newton methods. However, both (4.P) and its dual are nonsmooth in general. This motivates considering the concept of *alternating minimization envelope* (AME): this is a real-valued (as opposed to *extended* real-valued) exact merit function for the dual problem, and is precisely the augmented Lagrangian associated with (4.P) evaluated at the primal points computed by AMA. Under very general assumptions on (4.P), the AME is continuously differentiable around the set of dual solutions and even strictly twice differentiable there. As a consequence, the AME allows to extend classical, smooth unconstrained optimization algorithms to the solution of the dual problem to (4.P), which is nonsmooth in general. In this chapter we propose a dual line-search method, which uses the AME as merit function to compute the stepsizes. The convergence properties of the proposed algorithm greatly improve over (fast) AMA when fast-converging directions, computed by means of quasi-Newton formulas, are followed. Furthermore, we show that the AME is equivalent to the *forward-backward envelope* (FBE, see [113, 145, 150]) associated with the dual problem.

### 4.1.1 Contributions

In the present chapter we deal with the case where  $g$  in (4.P) is composed with a linear mapping. In this case, even though  $g$  may possess an efficiently computable proximal mapping,  $g \circ A$  in general does not. This motivates addressing the dual problem of (4.P) instead. The contributions and organization of the present chapter can be summarized as follows.

- We propose NAMA, the *Newton-type alternating minimization algorithm* (Section 4.2, Algorithm 6), a generalization of the alternating minimization algorithm that performs a line-search step over the AME: the proposed algorithm relies on the very same alternating minimization operations as AMA.
- We show that the AME is equivalent to the FBE of the dual problem (Section 4.3). This observation extends a classical result by Rockafellar, relating the Moreau envelope and the augmented Lagrangian, to our setting where an additional strongly convex term is present.
- We show that the proposed method enjoys global sublinear convergence under standard assumptions, and local linear convergence assuming *calmness* of the subdifferentials of the problem terms (Section 4.4).
- We analyze the first- and second-order properties of the AME, by linking them to generalized second-order properties of the primal functions  $f$  and  $g$  (Section 4.5).
- We show that the proposed method converges asymptotically superlinearly when the dual problem has a (unique) strong dual minimum, and the line-search directions are selected so as to satisfy the Dennis-Moré condition, as it is the case when quasi-Newton update formulas are adopted (Section 4.6). The effectiveness of our

approach is demonstrated by numerical simulations on linear MPC problems (Section 4.7).

Differently from the approaches in [145, 94, 136], NAMA does not require the gradient of the envelope function, therefore no second order information on the smooth term is needed: this would severely limit its applicability in the present setting where the dual problem is solved. Furthermore, with respect to the approaches of [94, 136], the algorithm presented here possesses strong global convergence properties which are not typical of classical line-search methods. Differently from [150], despite the fact that the selected directions may not be descent directions and the line search is performed on the envelope function, NAMA is a descent method for the dual objective: this allows to simplify the convergence analysis of the method, and to show the global sublinear convergence rate for the dual cost and the primal iterates.

## 4.2 Background and proposed algorithm

In problem (4.P),  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$ ,  $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$ ,  $A \in \mathbb{R}^{m \times n}$ . An alternative formulation for (4.P) is

$$\begin{aligned} & \underset{x \in \mathbb{R}^n, z \in \mathbb{R}^m}{\text{minimize}} && f(x) + g(z) \\ & \text{subject to} && Ax = z. \end{aligned} \tag{P'}$$

Therefore we can define the *augmented Lagrangian* associated with (4.P), denoted as

$$\mathcal{L}_\gamma(x, z, y) = f(x) + g(z) + \langle y, Ax - z \rangle + \frac{\gamma}{2} \|Ax - z\|^2,$$

where  $\gamma \geq 0$ . We indicate by  $\mathcal{L} \equiv \mathcal{L}_0$  the ordinary Lagrangian function.

Without further specifying it, throughout this chapter we will work under the following basic assumption.



**Assumption 4.1.** *The following hold:*

- (i)  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is strongly convex with modulus  $\mu_f > 0$ ;<sup>1</sup>
- (ii)  $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  is closed, proper, convex;
- (iii) either  $f$  and  $g$  are piecewise linear-quadratic and (4.P) is feasible, or (4.P) is strictly feasible: there exists  $x \in \text{relint}(\text{dom } f)$  such that  $Ax \in \text{relint}(\text{dom } g)$ .

**Remark 4.2.1.** The Fenchel dual problem is

$$\underset{y \in \mathbb{R}^m}{\text{minimize}} \quad \psi(y) = f^*(-A^\top y) + g^*(y). \quad (4.D)$$

Assumption 4.1 ensures that strong duality holds [133, Thm. 11.42] [132, Cor. 31.2.1], and the set of dual optimal points, which we will denote by  $Y_\star = \text{argmin } \psi$ , is nonempty. In case (4.P) is strictly feasible, then  $Y_\star$  is also compact [80, Thm. VII.2.3.2], and because of convexity the dual cost  $\psi$  has compact level sets.

Strong convexity of  $f$  implies that the solution  $x_\star$  to (4.P) is unique. Furthermore, the conjugate  $f^*$  is Lipschitz continuously differentiable with modulus  $\mu_f^{-1}$  [133, Th. 12.60].

Finally, the Moreau envelope  $(g^*)^\gamma$  is strictly continuous [133, Ex. 10.32] and has  $\gamma^{-1}$ -Lipschitz gradient

$$\nabla(g^*)^\gamma(y) = \gamma^{-1} (y - \text{prox}_{\gamma g^*}(y)), \quad (4.1)$$

as shown in [9, Prop. 12.29] (both the cited results apply since  $g^*$  is proper, closed and convex).  $\square$

Any primal-dual solution  $(x_\star, y_\star)$  of (4.P)-(4.D) satisfies the first-order

---

<sup>1</sup>A function  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  has convexity modulus  $c \geq 0$  if  $h - \frac{c}{2} \|\cdot\|^2$  is convex.

necessary conditions

$$-A^\top y_\star \in \partial f(x_\star) \quad (\Leftrightarrow \quad x_\star = \nabla f^*(-A^\top y_\star)) \quad (4.2a)$$

$$y_\star \in \partial g(Ax_\star) \quad (\Leftrightarrow \quad Ax_\star \in \partial g^*(y_\star)), \quad (4.2b)$$

which are also sufficient under [Assumption 4.1](#) [9, Thm. 19.1]. A natural way to tackle (4.P) is to solve (4.D) by means of forward-backward splitting (or proximal gradient method): starting from an initial dual point  $y^0 \in \mathbb{R}^m$ , iterate

$$y^{k+1} = T_\gamma(y^k) = \text{prox}_{\gamma g^*}(y^k + \gamma A \nabla f^*(-A^\top y^k)) \quad (4.3)$$

for some positive step-size parameter  $\gamma$ . If we define the associated *fixed-point residual*

$$R_\gamma(y) = \gamma^{-1}(y - T_\gamma(y)),$$

then dual optimality can be characterized as follows:

$$y_\star \in Y_\star \Leftrightarrow y_\star \in \text{fix } T_\gamma \Leftrightarrow y_\star \in \text{zer } R_\gamma \quad \forall \gamma > 0. \quad (4.4)$$

Iterations (4.3) are easily shown [11, Lem. 3.2] to be equivalent to the following scheme, the *alternating minimization algorithm* (AMA)

$$x^k = x(y^k) = \underset{x \in \mathbb{R}^n}{\text{argmin}} \{f(x) + \langle y^k, Ax \rangle\}, \quad (4.5a)$$

$$z^k = z_\gamma(y^k) = \underset{z \in \mathbb{R}^m}{\text{argmin}} \mathcal{L}_\gamma(x^k, z, y^k), \quad (4.5b)$$

$$y^{k+1} = y^k + \gamma(Ax^k - z^k). \quad (4.5c)$$

Note that step (4.5b) can be equivalently formulated as

$$z^k = \text{prox}_{\gamma^{-1}g}(\gamma^{-1}y^k + Ax(y^k)).$$

---

**Algorithm 6** NAMA (Newton-type Alternating Minimization Algorithm)

---

**Inputs:**  $y^0 \in \mathbb{R}^m$ ,  $\gamma \in (0, \mu_f / \|A\|^2)$ ,  $\beta \in (0, 1)$

**Initialize:**  $k = 0$

- 1:  $x^k = \operatorname{argmin}_x \{f(x) + \langle y^k, Ax \rangle\}$ ,  $z^k = \operatorname{argmin}_z \mathcal{L}_\gamma(x^k, z, y^k)$
- 2: Choose a direction  $d^k \in \mathbb{R}^m$
- 3: Find the largest  $\tau_k = \beta^{i_k}$ ,  $i_k \in \mathbb{N}$ , such that

$$\mathcal{L}_\gamma(\tilde{x}^k, \tilde{z}^k, \tilde{y}^k) \geq \mathcal{L}_\gamma(x^k, z^k, y^k), \quad (4.7)$$

where

$$\begin{aligned} \tilde{y}^k &= y^k + \tau_k d^k + \gamma(1 - \tau_k)(Ax^k - z^k) \\ \tilde{x}^k &= \operatorname{argmin}_x \{f(x) + \langle \tilde{y}^k, Ax \rangle\} \\ \tilde{z}^k &= \operatorname{argmin}_z \mathcal{L}_\gamma(\tilde{x}^k, z, \tilde{y}^k) \end{aligned}$$

- 4:  $y^{k+1} = \tilde{y}^k + \gamma(A\tilde{x}^k - \tilde{z}^k)$ ,  $k \leftarrow k + 1$ , go to step 1
- 

Using the notation of (4.5),  $T_\gamma$  and  $R_\gamma$  can be expressed as

$$T_\gamma(y) = y + \gamma(Ax(y) - z_\gamma(y)) \quad \text{and} \quad R_\gamma(y) = z_\gamma(y) - Ax(y). \quad (4.6)$$

Iterations (4.5) are guaranteed to converge, in the sense that  $x^k \rightarrow x_\star$ , provided that  $\gamma \in (0, 2\mu_f / \rho(A^\top A))$ , as follows from [154, Prop. 3]. Moreover, the dual cost in this case converges sublinearly to the optimum with global rate  $O(1/k)$ , and the extrapolation techniques introduced by Nesterov [104, 105, 106] allow to obtain accelerated versions of AMA with an optimal global rate  $O(1/k^2)$ , see [11].

### 4.2.1 Newton-type alternating minimization algorithm

The convergence speed of (fast) AMA is affected by ill-conditioning of the problem, as it is the case for all first-order methods. To accelerate convergence, we propose Algorithm 6. An overview of the algorithm is as follows:

- Algorithm 6 is composed by the very same operations as AMA: in

fact, only alternating minimization steps with respect to  $x$  and  $z$  are performed.

- Step 3 computes a new dual iterate  $\tilde{y}^k$ , by performing a line-search over the augmented Lagrangian associated with (4.P) evaluated at the alternating minimization primal points: we will see that this is equivalent to the forward-backward envelope function associated with the dual problem (4.D).
- The line-search is performed using a convex combination of the “nominal” residual direction  $\gamma(Ax^k - z^k)$  and an “arbitrary” direction  $d^k$ , to be selected so as to ensure fast asymptotic convergence. This novel choice of direction ensures that the line-search is feasible at every iteration (*i.e.*, condition (4.7) holds for a sufficiently small stepsize) despite the fact that  $d^k$  may not be a direction of descent, as we will see.
- Step 4 will allow us to obtain global convergence rates, and it comes at no cost since vectors  $\tilde{y}^k, \tilde{x}^k, \tilde{z}^k$  have already been computed in the line-search. In a sense, this step robustifies the algorithmic scheme.

By appropriately choosing  $d^k$ , the algorithm is able to greatly improve the convergence of (fast) AMA: we will prove that the algorithm converges with superlinear asymptotic rate when Newton-type directions are selected. For this reason we refer to Algorithm 6 as *Newton-type alternating minimization algorithm* (NAMA).

**Remark 4.2.2** (AMA as special case). If in Algorithm 6 one sets  $d^k = 0$  for all  $k \geq 0$ , then one can trivially set  $\tau_k = 1$ . In this case  $y^{k+1} = y^k + \gamma(Ax^k - z^k)$ , and Algorithm 6 reduces to AMA, cf. (4.5).  $\square$

**Remark 4.2.3** (General equality constrained problems). For any proper,

closed, convex  $h : \mathbb{R}^k \rightarrow \overline{\mathbb{R}}$  and  $B \in \mathbb{R}^{m \times k}$ , a problem of the form

$$\begin{aligned} & \underset{x \in \mathbb{R}^n, z \in \mathbb{R}^k}{\text{minimize}} && f(x) + h(z) \\ & \text{subject to} && Ax + Bz = b, \end{aligned} \tag{P''}$$

can be rewritten as (4.P) by letting

$$g(w) = (Bh)(b - w) = \inf_{z \in \mathbb{R}^k} \{h(z) \mid Bz = b - w\}. \tag{4.8}$$

Function  $(Bh)$  is called *image of  $h$  under  $B$* . Its conjugate is  $(Bh)^* = h^* \circ B^\top$ , see [133, Thm. 11.23(b)]. If we further assume  $\text{relint}(\text{dom } h^*) \cap \text{range}(B^\top) \neq \emptyset$ , then function  $(Bh)$  is proper, closed, convex, see [81, Thm. E.2.2.3], therefore function  $g$  in (4.8) satisfies [Assumption 4.1\(ii\)](#). Note that in this case the  $z$ - and  $y$ -update steps of AMA read

$$\begin{aligned} z^k &= \underset{z \in \mathbb{R}^k}{\text{argmin}} \left\{ g(z) + \langle y^k, Bz \rangle + \frac{\gamma}{2} \|Ax^k + Bz - b\|^2 \right\}, \\ y^{k+1} &= y^k + \gamma(Ax^k + Bz^k - b), \end{aligned}$$

and similar modifications allow to adapt NAMA to this more general setting. In light of these observations, what follows readily applies to problems of the form (P'').  $\square$

## 4.2.2 Quasi-Newton directions

There is freedom in selecting  $d^k$  in [Algorithm 6](#). To accelerate the convergence of the iterates, one possible choice is to compute fast converging directions for the system of nonlinear equations  $R_\gamma(y) = 0$  characterizing dual optimal points, cf. (4.4). Specifically, in [Algorithm 6](#) we can set

$$d^k = B_k^{-1}(Ax^k - z^k), \tag{4.9}$$

for a sequence of nonsingular matrices  $(B_k)_{k \in \mathbb{N}}$  approximating the inverse of the Jacobian  $JR_\gamma$  at the limit point of the dual iterates  $(y^k)_{k \in \mathbb{N}}$ . In quasi-Newton methods, starting from an initial nonsingular matrix  $B_0$ , the sequence of matrices  $(B_k)_{k \in \mathbb{N}}$  is determined by low-rank *updates* that satisfy the secant condition: in [Algorithm 6](#) fast asymptotic convergence can be proved if

$$B_{k+1}p^k = q^k \quad \text{with} \quad \begin{cases} p^k = \tilde{y}^k - y^k, \\ q^k = (\tilde{z}^k - A\tilde{x}^k) - (z^k - Ax^k), \end{cases}$$

as will be discussed in [Section 4.6](#). Note that all quantities required to compute the vectors  $p^k, q^k$  are available as by-product of the iterations.

In [\[122\]](#) the modified Broyden update is proposed, that prescribes rank-one updates of the form

$$(modified) \text{ Broyden} \quad B_{k+1} = B_k + \theta_k \frac{(q^k - B_k p^k)(p^k)^\top}{\|p^k\|^2}. \quad (4.10)$$

Here,  $(\theta_k)_{k \in \mathbb{N}} \subset [0, 2]$  is a sequence used to ensure that all terms in  $(B_k)_{k \in \mathbb{N}}$  are nonsingular, so that [\(4.9\)](#) is well defined. The original Broyden method [\[26\]](#) is obtained with  $\theta_k \equiv 1$ .

Probably the most popular quasi-Newton scheme is BFGS, which prescribes the following rank-two updates

$$BFGS \quad B_{k+1} = B_k + \frac{q^k(q^k)^\top}{\langle q^k, p^k \rangle} - \frac{B_k p^k (B_k p^k)^\top}{\langle p^k, B_k p^k \rangle}. \quad (4.11)$$

Note that in this case matrices  $B_k$  are symmetric, and in fact the fast asymptotic properties of BFGS are guaranteed only if the Jacobian  $JR_\gamma$  is symmetric [\[30\]](#) at the problem solution. This is not the case in our setting (cf. [Example 4.5.3](#)) although we have observed that [\(4.11\)](#) often outperforms other non-symmetric updates such as [\(4.10\)](#) in practice.

Using the Sherman-Morrison-Woodbury identity in [\(4.10\)](#) and [\(4.11\)](#) allows to directly store and update  $H_k = B_k^{-1}$ , so that  $d^k$  can be com-

puted without the need to invert matrices or solve linear systems.

Ultimately, instead of storing and operating on dense  $m \times m$  matrices, *limited-memory* variants of quasi-Newton schemes keep in memory only a few (usually 3 to 30) most recent pairs  $(p^k, q^k)$  implicitly representing the approximate inverse Jacobian. Their employment considerably reduces storage and computations over the full-memory counterparts, and as such they are the methods of choice for large-scale problems. The most popular limited-memory method is probably L-BFGS, which is based on the update (4.11), but efficiently computes matrix-vector products with the approximate inverse Jacobian using a *two-loop recursion* procedure [93, 107, 108].

### 4.3 Alternating minimization envelope

The fundamental tool enabling fast convergence of Algorithm 6 is the *alternating minimization envelope* function associated with (4.P). This is precisely the (negative) augmented Lagrangian function, evaluated at the primal points given by the alternating minimization steps.

**Definition 4.3.1.** *The alternating minimization envelope (AME) for (4.P), with parameter  $\gamma > 0$ , is the function*

$$\psi_\gamma(y) = -\mathcal{L}_\gamma(x(y), z_\gamma(y), y).$$

The first observation that we make relates the alternating minimization envelope in Definition 4.3.1 with the concept of *forward-backward envelope*.

**Theorem 4.3.2.** *Function  $\psi_\gamma$  is the forward-backward envelope (cf. [145, Def. 2.1]) associated with the dual problem (4.D):*

$$\begin{aligned} \psi_\gamma(y) = & f^*(-A^\top y) + g^*(T_\gamma(y)) + \frac{\gamma}{2} \|Ax(y) - z_\gamma(y)\|^2 \\ & + \gamma \langle Ax(y), z_\gamma(y) - Ax(y) \rangle, \end{aligned} \quad (4.12)$$

*Proof.* The optimality conditions for  $x(y)$  and  $z_\gamma(y)$  are

$$\partial f(x(y)) \ni -A^\top y,$$

$$\partial g(z_\gamma(y)) \ni T_\gamma(y) = y + \gamma(Ax(y) - z_\gamma(y)).$$

From these two, using (1.14), we obtain

$$f(x(y)) + f^*(-A^\top y) = -\langle Ax(y), y \rangle \quad (4.13a)$$

$$g(z_\gamma(y)) + g^*(T_\gamma(y)) = \langle z_\gamma(y), T_\gamma(y) \rangle \quad (4.13b)$$

Summing (4.13) and rearranging the terms we get (4.12).  $\square$

An alternative expression for  $\psi_\gamma$  in terms of the Moreau envelope of  $g^*$  is as follows, see [113]:

$$\psi_\gamma(y) = f^*(-A^\top y) - \frac{\gamma}{2} \|Ax(y)\|^2 + (g^*)^\gamma(y + \gamma Ax(y)). \quad (4.14)$$

The AME enjoys several favorable properties, some of which we now summarize. For any  $\gamma > 0$ ,  $\psi_\gamma$  is (strictly) continuous over  $\mathbb{R}^m$ , whereas if  $\gamma$  is small enough then the problem of minimizing  $\psi_\gamma$  is equivalent to solving (4.D). These properties are listed in the next result. In light of Theorem 4.3.2, this is equivalent to Proposition 2.2.2 and Proposition 2.2.3: here an alternative proof is shown, that exploits duality relations.

**Theorem 4.3.3.** *For any  $\gamma > 0$ ,  $\psi_\gamma$  is a strictly continuous function on  $\mathbb{R}^m$  satisfying*

$$(i) \quad \psi_\gamma(y) \leq \psi(y) + \frac{\gamma}{2} \|Ax(y) - z_\gamma(y)\|^2,$$

$$(ii) \quad \psi_\gamma(y) \geq \psi(T_\gamma(y)) + \frac{\gamma}{2} \left(1 - \frac{\gamma \|A\|^2}{\mu_f}\right) \|Ax(y) - z_\gamma(y)\|^2,$$

for any  $y \in \mathbb{R}^m$ . In particular, if  $\gamma < \mu_f / \|A\|^2$ , then the following also holds

$$(iii) \quad \inf \psi_\gamma = \inf \psi \text{ and } \operatorname{argmin} \psi_\gamma = \operatorname{argmin} \psi.$$

*Proof.* Strict continuity of  $\psi_\gamma$  follows immediately by the expression (4.14).



4.3.3(i): follows by [Lem. 4.A.1](#) using  $w = y$ .

4.3.3(ii): due to strong convexity of  $f$ ,  $f^*$  has  $1/\mu_f$ -Lipschitz gradient, and consequently

$$\begin{aligned} f^*(-A^\top T_\gamma(y)) &\leq f^*(-A^\top y) - \langle Ax(y), T_\gamma(y) - y \rangle + \frac{1}{2\mu_f} \|A^\top(T_\gamma(y) - y)\|^2 \\ &= f^*(-A^\top y) - \gamma \langle Ax(y), Ax(y) - z_\gamma(y) \rangle \\ &\quad + \frac{\gamma^2}{2\mu_f} \|A^\top(Ax(y) - z_\gamma(y))\|^2 \end{aligned} \quad (4.15)$$

Combining (4.12) with (4.15):

$$\begin{aligned} \psi_\gamma(y) &\geq \psi(T_\gamma(y)) - \frac{\gamma^2}{2\mu_f} \|A^\top(Ax(y) - z_\gamma(y))\|^2 + \frac{\gamma}{2} \|Ax(y) - z_\gamma(y)\|^2 \\ &\geq \psi(T_\gamma(y)) + \frac{\gamma}{2} \left(1 - \frac{\gamma \|A\|^2}{\mu_f}\right) \|Ax(y) - z_\gamma(y)\|^2. \end{aligned}$$

4.3.3(iii): easily follows combining [4.3.3\(i\)](#) and [4.3.3\(ii\)](#) with  $y = y_\star \in Y_\star$ , in light of the dual optimality condition (4.4).  $\square$

### 4.3.1 Analogy with the dual Moreau envelope

[Theorem 4.3.2](#) highlights a clear connection between the augmented Lagrangian, the forward-backward envelope and the alternating minimization algorithm. This closely resembles the one, first noticed by Rockafellar [[127](#), [128](#)], relating the augmented Lagrangian, the Moreau envelope and the *method of multipliers* (also known as *augmented Lagrangian method*) by Hestenes and Powell [[79](#), [121](#)]. Consider the general linear equality constrained convex problem

$$\begin{aligned} &\underset{z \in \mathbb{R}^k}{\text{minimize}} \quad g(z) \\ &\text{subject to} \quad Bz = b, \end{aligned} \quad (4.16)$$

where  $g : \mathbb{R}^m \rightarrow \overline{\mathbb{R}}$  is proper, closed, convex,  $B \in \mathbb{R}^{m \times k}$  and  $b \in \mathbb{R}^m$ . When applied to the dual of (4.16), namely

$$\underset{y \in \mathbb{R}^m}{\text{minimize}} \quad \omega(y) = g^*(-B^\top y) + \langle b, y \rangle,$$

the proximal minimization algorithm [18, §5.2] is equivalent to the following augmented Lagrangian method

$$\begin{aligned} z^k &= \underset{z \in \mathbb{R}^n}{\text{argmin}} \{g(z) + \langle y^k, Bz - b \rangle + \frac{\gamma}{2} \|Bz - b\|^2\}, \\ y^{k+1} &= y^k + \gamma(Bz^k - b). \end{aligned}$$

If  $\text{range}(B^\top) \cap \text{relint}(\text{dom } g^*) \neq \emptyset$  one can show, with a similar proof to that of Theorem 4.3.2, that the Moreau envelope of  $\omega$  satisfies

$$\begin{aligned} \omega^\gamma(y^k) &= -g(z^k) - \langle y^k, Bz^k - b \rangle - \frac{\gamma}{2} \|Bz^k - b\|^2 \\ &= -\mathcal{L}_\gamma(z^k, y^k). \end{aligned}$$

Therefore the forward-backward and Moreau envelope functions have the same nice interpretation in terms of augmented Lagrangian, when they are applied to the dual of equality constrained convex problems: in a sense, Theorem 4.3.2 extends and generalizes the classical result on the dual Moreau envelope, by allowing for an additional variable  $x$  and a strongly convex term  $f$  in the problem.

## 4.4 Convergence

We now turn our attention to the global convergence properties of Algorithm 6.

**Remark 4.4.1** (Feasibility of the line-search). The line-search step 3 is well defined regardless of the choice of  $d^k$ : at any iteration  $k$  condition (4.7) holds for  $i_k$  sufficiently large. To see this, suppose that  $\|Ax^k - z^k\| > 0$  (otherwise  $(x^k, y^k)$  is a primal-dual solution). Then, since  $\gamma < \mu_f/\|A\|^2$ ,

Theorem 4.3.3 implies that

$$\psi_\gamma(T_\gamma(y^k)) < \psi_\gamma(y^k). \quad (4.17)$$

Since  $\tilde{y}^k \rightarrow T_\gamma(y^k)$  as  $\tau_k \rightarrow 0$  and  $\psi_\gamma$  is continuous, then necessarily  $\psi_\gamma(\tilde{y}^k) \leq \psi_\gamma(y^k)$  for  $\tau_k$  sufficiently small.  $\square$

Theorem 4.3.3 also ensures that the following chain of inequalities, which will be fundamental for the following convergence results, holds in Algorithm 6:

$$\psi(y^{k+1}) \leq \psi_\gamma(\tilde{y}^k) \quad (4.18a)$$

$$\leq \psi_\gamma(y^k) \quad (4.18b)$$

$$\leq \psi(y^k) - \frac{\gamma}{2} \|Ax^k - z^k\|^2. \quad (4.18c)$$

In particular, Algorithm 6 is a descent method for  $\psi$ .

**Remark 4.4.2** (Backtracking on  $\gamma$ ). In practice, no prior knowledge of the global Lipschitz constant  $\|A\|^2/\mu_f$  is required for Algorithm 6: instead of using a fixed parameter  $\gamma$ , one can determine a non-increasing sequence of parameters  $(\gamma_k)_{k \in \mathbb{N}}$  ensuring that the inequalities (4.17) and (4.18a) still hold at every iteration. This is analogous to what is done in practice in (fast) AMA, see [11, Rem. 3.4] and [10, §3, §4]. If such a strategy is implemented, all the convergence results that follow will either remain unchanged, or continue to hold provided that eventually  $\gamma_k < \mu_f/\|A\|^2$ .  $\square$

We now prove that the iterates of (6) converge to the dual optimal cost and to the primal solution. Moreover, global convergence rates are provided.

**Theorem 4.4.3** (Global convergence). *In Algorithm 6:*

- (i)  $Ax^k - z^k \rightarrow 0$  and all cluster points of  $(y^k)_{k \in \mathbb{N}}$  are dual optimal, i.e., they belong to  $Y_\star$ ;

(ii) if (4.P) is strictly feasible then  $\psi(y^k) \searrow \inf \psi$  with global rate  $O(1/k)$ , and  $x^k \rightarrow x_*$  with global rate  $O(1/\sqrt{k})$ ;

(iii) if both  $f$  and  $g$  are piecewise linear-quadratic then  $\psi(y^k) \searrow \inf \psi$  with global  $Q$ -linear rate, and  $x^k \rightarrow x_*$  with global  $R$ -linear rate.

*Proof.* 4.4.3(i): by (4.18c), for all  $i \geq 0$  we have

$$\psi(y^{i+1}) \leq \psi(y^i) - \frac{\gamma}{2} \|Ax^i - z^i\|^2.$$

Summing up the inequality for  $i = 1, \dots, k$  we obtain

$$\inf \psi \leq \psi(y^{k+1}) \leq \psi(y^1) - \frac{\gamma}{2} \sum_{i=1}^k \|Ax^i - z^i\|^2$$

(the sum starts from  $i = 1$  since  $y^0$  may be dual infeasible). In particular (cf. (4.6))  $R_\gamma(y^k) = z^k - Ax^k \rightarrow 0$ , and since  $R_\gamma$  is continuous, necessarily all cluster points of  $(y^k)_{k \in \mathbb{N}}$  are optimal.

4.4.3(ii): the proof is similar to that of [106, Thm. 4]. Let  $D > 0$  be such that  $\text{dist}(y, Y_*) < D$  for all points  $y \in \{y \in \mathbb{R}^m \mid \psi(y) \leq \psi(y^0)\}$ . Such a constant  $D$  exists because  $\psi$  has bounded level sets (cf. Rem. 4.2.1). Moreover, from [145, Prop. 2.5] we know that  $\psi_\gamma \leq \psi^\gamma$  (the Moreau envelope of  $\psi$ ). Therefore,

$$\begin{aligned} \psi(y^{k+1}) &\stackrel{(4.18b)}{\leq} \psi_\gamma(y^k) \\ &\leq \psi^\gamma(y^k) = \min_{w \in \mathbb{R}^m} \left\{ \psi(w) + \frac{1}{2\gamma} \|w - y^k\|^2 \right\}. \end{aligned}$$

In particular, for  $y_* \in \text{argmin } \psi$ ,

$$\begin{aligned} \psi(y^{k+1}) &\leq \min_{\alpha \in [0,1]} \left\{ \psi(\alpha y_* + (1-\alpha)y^k) + \frac{\alpha^2}{2\gamma} \|y^k - y_*\|^2 \right\} \\ &\leq \min_{\alpha \in [0,1]} \left\{ \psi(y^k) - \alpha(\psi(y^k) - \inf \psi) + \frac{D^2}{2\gamma} \alpha^2 \right\} \end{aligned}$$

where in last inequality we used convexity of  $\psi$ . In case  $\psi(y^0) - \inf \psi \geq$

$D^2/\gamma$ , then the optimal solution of the latter problem for  $k = 0$  is  $\alpha = 1$ , and  $\psi(y^1) - \inf \psi \leq D^2/2\gamma$ . Otherwise, the optimal solution is

$$\alpha = \gamma \frac{\psi(y^k) - \inf \psi}{D^2} \leq \gamma \frac{\psi(y^0) - \inf \psi}{D^2} \leq 1$$

and we obtain

$$\psi(y^{k+1}) \leq \psi(y^k) - \gamma \frac{(\psi(y^k) - \inf \psi)^2}{2D^2}.$$

Letting  $\lambda_k = \frac{1}{\psi(y^k) - \inf \psi}$  the latter inequality is expressed as

$$\frac{1}{\lambda_{k+1}} \leq \frac{1}{\lambda_k} - \frac{\gamma}{2D^2 \lambda_{k+1}^2}.$$

Multiplying both sides by  $\lambda_k \lambda_{k+1}$  and rearranging,

$$\lambda_{k+1} \geq \lambda_k + \frac{\gamma}{2D^2} \frac{\lambda_{k+1}}{\lambda_k} \geq \lambda_k + \frac{\gamma}{2D^2},$$

where the latter inequality follows from the fact that  $(\psi(y^k))_{k \in \mathbb{N}}$  is non-increasing, as shown in (4.18). Telescoping the inequality we obtain

$$\lambda_k \geq \lambda_0 + k \frac{\gamma}{2D^2} \geq k \frac{\gamma}{2D^2},$$

and therefore  $\psi(y^k) - \inf \psi \leq 2D^2/k\gamma$ . This, together with [Lem. 4.A.4](#), proves [4.4.3\(ii\)](#).

[4.4.3\(iii\)](#): using (4.18) we have that

$$\psi(y^k) - \psi(y^{k+1}) \geq \frac{\gamma}{2} \|Ax^k - z^k\|^2. \quad (4.19)$$

Furthermore, using [Lem. 4.A.1](#) with  $w = y_* = \Pi_{Y_*} y^k$  and  $y = y^k$ , we

obtain

$$\begin{aligned}\psi(y^{k+1}) - \inf \psi &\leq \psi_\gamma(y^k) - \inf \psi \\ &\leq \langle Ax^k - z^k, y_\star - y^k \rangle - \frac{\gamma}{2} \|Ax^k - z^k\|^2,\end{aligned}$$

where first inequality is due to (4.18c). This implies

$$\psi(y^{k+1}) - \inf \psi \leq \|Ax^k - z^k\|^2 \left( \frac{\text{dist}(y^k, Y_\star)}{\|Ax^k - z^k\|} - \frac{\gamma}{2} \right)$$

which, using (4.19), yields

$$\psi(y^{k+1}) - \inf \psi \leq \left( 1 - \frac{\gamma}{2} \frac{\|Ax^k - z^k\|}{\text{dist}(y^k, Y_\star)} \right) (\psi(y^k) - \inf \psi). \quad (4.20)$$

Functions  $f^*$  and  $g^*$  are piecewise linear-quadratic in this case [133, Thm. 11.14]. Therefore by [89, Thm. 2.7]  $\psi$  enjoys the following quadratic growth condition globally: for any  $\nu > 0$  there is  $\alpha > 0$  such that

$$\frac{\alpha}{2} \text{dist}^2(y, Y_\star) \leq \psi(y^k) - \inf \psi \quad \forall y : \psi(y) - \inf \psi \leq \nu,$$

which by [54, Cor. 3.6] is equivalent to the following error bound condition for some  $\beta > 0$

$$\text{dist}(y, Y_\star) \leq \beta \|Ax(y) - z_\gamma(y)\| \quad \forall y : \psi(y) - \inf \psi \leq \nu, \quad (4.21)$$

Using (4.21) in (4.20) we obtain global Q-linear convergence of  $(\psi(y^k))_{k \in \mathbb{N}}$ , and by Lem. 4.A.4 we obtain R-linear convergence of  $(y^k)_{k \in \mathbb{N}}$ .  $\square$

When  $f$  and  $g$  are not both piecewise linear-quadratic, we can prove local linear convergence of Algorithm 6 provided that their subdifferentials are *calm* at  $x_\star$ , see Definition 1.A.3.

Calmness is a very common property of the subdifferential mapping. The subdifferential of all piecewise linear-quadratic functions is calm everywhere, as follows from [52, Prop. 3H.1]. Other examples include the nuclear and spectral norms [139]. Smooth functions, *i.e.*, with Lip-

schild gradient, clearly have calm subdifferential: this includes Moreau envelopes of closed, convex functions, such as the Huber loss for robust estimation, and commonly used loss functions such as the squared Euclidean norm and the logistic loss.

Calmness is equivalent to metric subregularity of the inverse mapping [52, Thm. 3H.3]: from [163, Prop. 6, Prop. 8] we then deduce that the indicator functions of  $\ell_1$ ,  $\ell_\infty$  and Euclidean norm balls all have calm subdifferentials.

The following result holds. Its proof is analogous to the one of [54, Thm. 4.2], although our assumption of calmness is equivalent to metric subregularity of  $\partial f^*$  and  $\partial g^*$ , which is weaker than the firm convexity assumed in [54].

**Theorem 4.4.4** (Local linear convergence). *Suppose that strict complementarity holds in (4.P), that  $\partial f$  is calm at  $x_*$  and  $\partial g$  is calm at  $Ax_*$ . Then in Algorithm 6 eventually  $\psi(y^k) \rightarrow \inf \psi$  with  $Q$ -linear rate and  $x^k \rightarrow x_*$  with  $R$ -linear rate.*

*Proof.* As discussed in the proof of Thm. 4.4.3(iii), it suffices to show that an error bound of the form (4.21) holds for some  $\beta, \nu > 0$ .

The assumed calmness properties of  $\partial f$  and  $\partial g$  are equivalent to *metric subregularity* of  $\partial f^*$  at  $-A^\top y_*$  for  $x_*$ , and of  $\partial g^*$  at  $y_*$  for  $Ax_*$ , see [52, Thm. 3H.3], for all  $y_* \in Y_*$ . This can be seen, using [2, Thm. 3.3], to be equivalent to the following: there is  $c_{y_*} > 0$  and a neighborhood  $U_{y_*}$  of  $y_*$  such that for all  $y \in U_{y_*}$

$$\begin{aligned} f^*(-A^\top y) &\geq f^*(-A^\top y_*) + \langle x_*, A^\top (y_* - y) \rangle \\ &\quad + \frac{c_{y_*}}{2} \text{dist}^2(-A^\top y, (\nabla f^*)^{-1}(x_*)), \\ g^*(y) &\geq g^*(y_*) + \langle Ax_*, y - y_* \rangle + \frac{c_{y_*}}{2} \text{dist}^2(y, (\partial g^*)^{-1}(Ax_*)). \end{aligned}$$

Since  $Y_* \subset \bigcup_{y_* \in Y_*} U_{y_*}$  and  $Y_*$  is compact, we may select a *finite* subset  $W \subset Y_*$  such that  $Y_* \subset U_{Y_*} = \bigcup_{y_* \in W} U_{y_*}$ . Summing the above inequali-

ties for all  $y_\star \in W$ , and denoting  $c = \min \{c_{y_\star} \mid y_\star \in W\} > 0$ , we obtain

$$\psi(y) \geq \inf \psi + \frac{c}{2} [\text{dist}^2(-A^\top y, \partial f(x_\star)) + \text{dist}^2(y, \partial g(Ax_\star))] , \quad (4.22)$$

for all  $y \in U_{Y_\star}$ , where we have also used  $(\nabla f^\star)^{-1} = \partial f$  and  $(\partial g^\star)^{-1} = \partial g$ . Using [Lem. 4.A.5](#), and the fact that  $a^2 + b^2 \geq 2ab$  for any  $a, b \in \mathbb{R}$ , we obtain from [\(4.22\)](#) that for some  $\kappa > 0$

$$\psi(y) \geq \inf \psi + \frac{\kappa}{2} \text{dist}^2(y, Y_\star), \quad \forall y \in U_{Y_\star},$$

i.e.,  $\psi$  satisfies the quadratic growth condition, which by [\[54, Cor. 3.6\]](#) is equivalent to the error bound condition [\(4.21\)](#). This completes the proof.  $\square$

## 4.5 First- and second-order properties

[Algorithm 6](#) is a line-search method for the unconstrained minimization of  $\psi_\gamma$  which is, by [Theorem 4.3.3\(iii\)](#), equivalent to solving [\(4.D\)](#). To enable fast convergence of the iterates, we can apply ideas from smooth unconstrained optimization in selecting the sequence  $(d^k)_{k \in \mathbb{N}}$  of directions. To this end, differentiability of  $\psi_\gamma$  around dual solutions  $y_\star$  is a desirable property. We will now see that this is implied by generalized second-order properties of  $f$  around  $x_\star$ , which are introduced in the following assumption. Analogous assumptions on  $g$  further ensure that  $\psi_\gamma$  is (strictly) twice differentiable at  $y_\star$ . The interested reader is referred to [\[133\]](#) for an extensive discussion on (second-order) epi-differentiability.

**Assumption 4.2.** *With respect to a primal-dual solution  $(x_\star, y_\star)$ ,*

- (i) *function  $f$  is strictly twice epi-differentiable at all  $x \in \text{dom } f$  close enough to  $x_\star$ , and the second-order epi-derivative at  $x_\star$  for  $-A^\top y_\star$  is in particular*

$$\text{d}^2 f(x_\star | -A^\top y_\star)[w] = \langle H_f w, w \rangle + \delta_{S_f}(w), \quad \forall w \in \mathbb{R}^m, \quad (4.23)$$



where  $S_f$  is a linear subspace of  $\mathbb{R}^m$  and  $H_f \in \mathbb{R}^{m \times m}$ ;

(ii)  $g$  is (strictly) twice epi-differentiable at  $Ax_*$  for  $y_*$ , with

$$d^2g(Ax_*|y_*)[w] = \langle H_g w, w \rangle + \delta_{S_g}(w), \quad \forall w \in \mathbb{R}^m, \quad (4.24)$$

where  $S_g$  is a linear subspace and  $H_g \in \mathbb{R}^{n \times n}$ .

When the stronger condition in parenthesis is required, the assumptions will be referred to as [Assumption 4.2<sup>+</sup>](#).

Whenever [Assumption 4.2](#) holds, we can consider  $H_f$  symmetric and positive semidefinite, satisfying  $\text{range}(H_f) = S_f$  and  $\text{null}(H_f) = S_f^\perp$ . Similarly,  $H_g$  can be assumed to be symmetric and positive semidefinite satisfying  $\text{range}(H_g) \subseteq S_g$  and  $\text{null}(H_g) \supseteq S_g^\perp$ . This causes no loss of generality: matrix  $H'_f = \frac{1}{2} \Pi_{S_f} (H_f + H_f^\top) \Pi_{S_f}$  has the desired properties and satisfies (4.23) provided  $H_f$  does, and similarly for  $H_g$ .

**Theorem 4.5.1** (Differentiability of  $\psi_\gamma$ ). *Suppose that [Assumption 4.2\(i\)](#) holds for a primal-dual solution  $(x_*, y_*)$ . Then  $\psi_\gamma$  is of class  $C^1$  around  $y_*$ , with*

$$\nabla \psi_\gamma(y) = Q_\gamma(y) R_\gamma(y)$$

where  $Q_\gamma(y) = I - \gamma A \nabla^2 f^*(-A^\top y) A^\top$ .

*Proof.* From [Lem. 4.A.2](#) it follows that  $\hat{f} = f^* \circ (-A^\top)$  is of class  $C^2$  around  $y_*$ . The claim now easily follows from the chain rule of differentiation applied to (4.14), by using (4.1).  $\square$

Twice differentiability of  $\psi_\gamma$  at  $y_*$  is very important: when Newton-type directions are used, this implies that eventually unit stepsize will be accepted and fast asymptotic convergence will take place. In other words, unlike standard nonsmooth merit functions for constrained optimization,  $\psi_\gamma$  does not prevent the acceptance of unit stepsize.

**Theorem 4.5.2** (Twice differentiability of  $\psi_\gamma$ ). *Suppose that [Assumption 4.2](#) holds with respect to a primal-dual solution  $(x_*, y_*)$ . Then,*

(i)  $R_\gamma$  is differentiable at  $y_\star$  with Jacobian

$$JR_\gamma(y_\star) = \gamma^{-1} [I - P_\gamma(y_\star)Q_\gamma(y_\star)]; \quad (4.25)$$

here,  $Q_\gamma$  is as in [Theorem 4.5.1](#) and

$$P_\gamma(y_\star) = J \operatorname{prox}_{\gamma g^*} (y_\star + \gamma A \nabla f^*(-A^\top y_\star)) = \Pi_{\bar{S}} (I + \gamma \bar{G})^{-1} \Pi_{\bar{S}} \quad (4.26)$$

with  $\bar{G} = \Pi_{S_g} H_g^\dagger \Pi_{S_g}$  and  $\bar{S} = S_g^\perp + \operatorname{range}(H_g)$ ;

(ii)  $\psi_\gamma$  is twice differentiable at  $y_\star$  with symmetric Hessian

$$\nabla^2 \psi_\gamma(y_\star) = \gamma^{-1} Q_\gamma(y_\star) [I - P_\gamma(y_\star)Q_\gamma(y_\star)]. \quad (4.27)$$

Moreover, if [Assumption 4.2<sup>+</sup>](#) holds then  $R_\gamma$  and  $\nabla \psi_\gamma$  are strictly differentiable at  $y_\star$ .

*Proof.* The proof proceeds similarly to the ones of [Lem. 2.2.9](#) and [Thm. 2.2.10](#) by using duality properties, provided in [§4.A](#), of second order epi-derivatives.

Let  $\hat{f} = f^* \circ (-A^\top)$  and  $L_{\hat{f}} = \mu_f / \|A\|^2$ . We know from [\[120, Thms. 3.8, 4.1\]](#) and [\[133, Thm. 13.21\]](#) that  $\operatorname{prox}_{\gamma g^*}$  is (strictly) differentiable at  $y_\star - \gamma \nabla \hat{f}(y_\star)$  if and only if  $g$  satisfies [Ass. 4.2\(ii\)](#) ([4.2\(ii\)<sup>+</sup>](#)); in fact, by [\(4.2\)](#) we know that  $Ax_\star = -\nabla \hat{f}(y_\star)$ . Moreover, due to [Lem. 4.A.2](#),  $\hat{f} \in C^2$  in a neighborhood of  $y_\star$  and in particular  $\nabla \hat{f}$  is strictly differentiable at  $y_\star$ . The formula for  $JR_\gamma(y_\star)$  follows from [\(4.1\)](#) and the chain rule of differentiation.

We now prove the claimed expression for  $P_\gamma(y_\star)$ . We may invoke [Lem. 4.A.3](#) and apply [\[133, Ex. 13.45\]](#) to the *tilted* function  $g + \langle \nabla \hat{f}(y_\star), \cdot \rangle$ :

this tells us that for all  $d \in \mathbb{R}^m$

$$\begin{aligned}
P_\gamma(y_\star)d &= \text{prox}_{(\gamma/2)d^2g^*(y_\star|Ax_\star)}(d) \\
&= \operatorname{argmin}_{d' \in \bar{S}} \left\{ \frac{1}{2} \langle d', \bar{G}d' \rangle + \frac{1}{2\gamma} \|d' - d\|^2 \right\} \\
&= \Pi_{\bar{S}} \operatorname{argmin}_{d' \in \mathbb{R}^n} \left\{ \frac{1}{2} \langle \Pi_{\bar{S}} d', \bar{G} \Pi_{\bar{S}} d' \rangle + \frac{1}{2\gamma} \|\Pi_{\bar{S}} d' - d\|^2 \right\} \\
&= \Pi_{\bar{S}} (\Pi_{\bar{S}} [I + \gamma \bar{G}] \Pi_{\bar{S}})^\dagger \Pi_{\bar{S}} d \\
&= \Pi_{\bar{S}} [I + \gamma \bar{G}]^{-1} \Pi_{\bar{S}} d
\end{aligned}$$

where  $^\dagger$  indicates the pseudo-inverse, and last equality is due to [14, Facts 6.4.12(i)-(ii) and 6.1.6(xxxii)]. In fact, letting  $U = \Pi_{\bar{S}}$  and  $V = \Pi_{\bar{S}}[I + \gamma \bar{G}]$  we have that  $\text{range}(U^\top UV) = \text{range}(V)$  and  $\text{range}(VV^\top U^\top) \subseteq \bar{S} = \text{range}(U)$ , as required by [14, Facts 6.4.12(ii)].

With basic calculus rules one can verify that, since  $R_\gamma(y_\star) = 0$ ,  $\nabla \psi_\gamma = Q_\gamma R_\gamma$  is (strictly) differentiable at  $y_\star$  provided that  $Q_\gamma$  is (strictly) continuous at  $y_\star$  and  $R_\gamma$  is (strictly) differentiable at  $y_\star$ . A simple application of the chain rule of differentiation concludes the proof of 4.5.2(ii).  $\square$

To better understand the requirements of [Assumption 4.2](#), let us consider the following simple but significant example: when  $f$  is  $C^2$  and  $g \circ A$  models linear inequality constraints, [Assumption 4.2](#) is implied by *strict complementarity*.

**Example 4.5.3** ( $C^2$  functions subject to polyhedral constraints). Consider problems of the form

$$\underset{x \in \mathbb{R}^n}{\text{minimize}} \quad f(x) + \delta_C(Ax),$$

where  $g = \delta_C$  is the indicator of  $C = \{z \in \mathbb{R}^m \mid z \leq b\}$ ,  $b \in \mathbb{R}^m$ , and  $f \in C^2$ . In this case [Assumption 4.2\(i\)](#) holds with  $H_f = \nabla^2 f(x_\star)$ ,  $S_f = \mathbb{R}^n$  (therefore  $\Pi_{S_f} = \text{Id}$ ), see [133, Ex. 13.8]. Regarding [Assumption 4.2\(ii\)](#),

one can use [133, Ex. 13.17] to see that

$$d^2g(Ax_\star|y_\star)[w] = \delta_{K(Ax_\star, y_\star)}(w),$$

where  $K$  is the critical cone. Denoting as  $T_C(y)$  the tangent cone of set  $C$  at  $y \in C$ , and as  $J = \{i \mid (Ax_\star)_i = b_i\}$  the set of active constraints at the solution  $x_\star$ , the critical cone is given by

$$\begin{aligned} K(Ax_\star, y_\star) &= \{w \in T_C(Ax_\star) \mid \langle y_\star, w \rangle = 0\}, \\ &= \{w \mid \langle y_\star, w \rangle = 0, w_i \leq 0 \ \forall i \in J\}. \end{aligned}$$

For  $K(Ax_\star, y_\star)$  to be a subspace, necessarily  $(y_\star)_i > 0$  for all  $i \in J$ , *i.e.*, strict complementarity must hold at the primal-dual solution  $(x_\star, y_\star)$ . In this case, [Assumption 4.2\(ii\)](#) holds with  $H_g = 0$  and

$$S_g = K(Ax_\star, y_\star) = \{w \mid w_i = 0 \ \forall i \in J\}.$$

We may assume that  $J = \{1, \dots, k\}$  without loss of generality, *i.e.*, the first  $k$  constraints are the active ones, and let  $\bar{J} = \{1, \dots, m\} \setminus J$ . Note that  $\nabla^2 f^*(-A^\top y_\star) = \nabla^2 f(x_\star)^{-1}$  due to strong convexity of  $f$ , see [133, Ex. 11.9]. Partitioning the inverse Hessian and constraint matrix as

$$\nabla^2 f(x_\star)^{-1} = \begin{bmatrix} H_{JJ} & H_{J\bar{J}} \\ H_{\bar{J}J} & H_{\bar{J}\bar{J}} \end{bmatrix}, \quad A = \begin{bmatrix} A_J \\ A_{\bar{J}} \end{bmatrix},$$

and using the notation of [Theorem 4.5.2\(i\)](#), by elementary computations one sees that

$$P_\gamma(y_\star) = \begin{bmatrix} I_k & 0 \\ 0 & 0 \end{bmatrix}, \quad JR_\gamma(y_\star) = \begin{bmatrix} A_J H_{JJ} A_J^\top & A_J H_{J\bar{J}} A_{\bar{J}}^\top \\ 0 & \gamma^{-1} I_{m-k} \end{bmatrix}. \quad (4.28)$$

□

Finally, we can relate strong minimality of  $\psi$  and  $\psi_\gamma$  to nonsingularity of the Jacobian of  $R_\gamma$  and to the generalized second-order properties of  $f$

and  $g$  as follows.

**Theorem 4.5.4** (Conditions for strong minimality). *If [Assumption 4.2](#) holds for a primal-dual solution  $(x_*, y_*)$ , then for all  $\gamma < \mu_f / \|A\|^2$  the following are equivalent:*

- (a)  $y_*$  is a strong minimum for  $\psi$ ;
- (b)  $JR_\gamma(y_*)$  is nonsingular;
- (c) the (symmetric) matrix  $\nabla^2 \psi_\gamma(y_*)$  is positive definite;
- (d)  $y_*$  is a strong minimum for  $\psi_\gamma$ .

*Proof.* The proof proceeds similarly to the one of [Thm. 2.2.11](#) by using duality properties, provided in [§4.A](#), of second order epi-derivatives.

Let  $\bar{G}$  and  $\bar{S}$  be as in [Thm. 4.5.2\(i\)](#), and  $H_f$  and  $S_f$  as in [Ass. 4.2](#). From [Lem. 4.A.2](#) and [Thm. 4.5.2\(ii\)](#) we know that  $\hat{f} = f^* \circ (-A^\top)$  is of class  $C^2$  around  $y_*$  with  $\nabla^2 \hat{f}(y) = A \Pi_{S_f} H_f^\dagger \Pi_{S_f} A^\top$ , and that  $\nabla^2 \psi_\gamma(y_*)$  exists and is symmetric. Moreover, it follows from [Lem. 4.A.3](#) and [[133](#), Ex. 13.18] that for all  $d \in \mathbb{R}^n$

$$\begin{aligned} d^2 \psi(y|0)[d] &= \langle d, \nabla^2 \hat{f}(y) d \rangle + d^2 g^*(y | -\nabla \hat{f}(y)) [d] \\ &= \langle d, A \Pi_{S_f} H_f^\dagger \Pi_{S_f} A^\top d \rangle + \langle d, \bar{G} d \rangle + \delta_{\bar{S}}(d). \end{aligned} \quad (4.29)$$

We will show that all the conditions are equivalent to

$$(e) \quad \langle Hd, d \rangle > 0 \quad \text{for all } d \in \bar{S}, \text{ where } H = A \Pi_{S_f} H_f^\dagger \Pi_{S_f} A^\top + \bar{G}.$$

[4.5.4\(a\)](#)  $\Leftrightarrow$  [4.5.4\(e\)](#): Follows from (4.29), using [[133](#), Thm. 13.24(c)].

[4.5.4\(b\)](#)  $\Leftrightarrow$  [4.5.4\(c\)](#): As it is apparent from [Rem. 4.2.1](#),  $Q_\gamma$  is invertible since  $\gamma < \mu_f / \|A\|^2$ . Therefore, from (4.25) and (4.27) we see that  $JR_\gamma(y_*)$  is similar to the symmetric matrix  $Q_\gamma(y_*)^{-1/2} \nabla^2 \psi_\gamma(y_*) Q_\gamma(y_*)^{-1/2}$ , which is positive definite, and in particular invertible, if and only if  $\nabla^2 \psi_\gamma(y_*)$  is.

4.5.4(e)  $\Leftrightarrow$  4.5.4(b): In the proof of the implication above, we showed that  $JR_\gamma(y_\star)$  is similar to a symmetric matrix, and in particular the condition  $\lambda_{\min}(JR_\gamma(y_\star)) > 0$  is sufficient for ensuring its nonsingularity. Let  $P = P_\gamma(y_\star)$  and  $Q = Q_\gamma(y_\star)$ , so that  $JR_\gamma(y_\star) = \gamma^{-1}(I - QP)$ . From [83, Thm. 7.7.3(a)] it follows that  $\lambda_{\min}(I - QP) > 0$  if and only if  $Q^{-1} \succ P$ . For all  $d \in \bar{S}$ , using (4.26) we have

$$\begin{aligned} \langle d, (Q^{-1} - P)d \rangle &= \langle d, Q^{-1}d \rangle - \langle d, \Pi_{\bar{S}}[I + \gamma\bar{G}]^{-1} \Pi_{\bar{S}} d \rangle \\ &= \langle d, Q^{-1}d \rangle - \langle \Pi_{\bar{S}} d, [I + \gamma\bar{G}]^{-1} \Pi_{\bar{S}} d \rangle \\ &= \langle d, Q^{-1}d \rangle - \langle d, [I + \gamma\bar{G}]^{-1}d \rangle \\ &= \langle d, Q^{-1}d \rangle - \langle d, [I + \gamma\bar{G}]^{-1}d \rangle \end{aligned}$$

and last quantity is positive if and only if  $I + \gamma\bar{G} \succ Q$  on  $\bar{S}$ . By definition of  $Q$ , we then have that this holds if and only if  $\nabla^2 \hat{f}(x) + \bar{G} = A \Pi_{S_f} H_f^\dagger \Pi_{S_f} A^\top + \bar{G} \succ 0$  on  $\bar{S}$ , which is 4.5.4(e). For  $d \in \bar{S}^\perp$  the computation trivializes to  $\langle d, (Q^{-1} - P)d \rangle = \langle d, Q^{-1}d \rangle > 0$  regardless.

4.5.4(c)  $\Leftrightarrow$  4.5.4(d): Trivial since  $\nabla^2 \psi_\gamma(y_\star)$  exists.  $\square$

In the context of Example 4.5.3, from (4.28) one has

$$JR_\gamma(y_\star) \text{ is nonsingular} \Leftrightarrow A_J H_{JJ} A_J^\top \text{ is nonsingular.}$$

Since  $\nabla^2 f(x_\star) \succ 0$  by assumption, then  $H_{JJ} \succ 0$  and nonsingularity of the Jacobian is equivalent to  $A_J$  being full row rank, *i.e.*, linear independence of the active constraints at  $x_\star$  (the LICQ assumption).

## 4.6 Superlinear convergence

The following definition (cf. [59, Eq. (7.5.2)]) gives the fundamental condition, on the sequence  $(d^k)_{k \in \mathbb{N}}$  of directions, ensuring superlinear asymptotic convergence of Algorithm 6.

**Definition 4.6.1.** For  $(y^k)_{k \in \mathbb{N}}$  converging to  $y_*$ , we say that  $(d^k)_{k \in \mathbb{N}}$  is super-linearly convergent w.r.t.  $(y^k)_{k \in \mathbb{N}}$  if

$$\lim_{k \rightarrow \infty} \frac{\|y^k + d^k - y_*\|}{\|y^k - y_*\|} = 0. \quad (4.30)$$

When  $y_*$  is a strong minimizer, by [54, Cor. 3.6] the error bound (4.21) holds for some  $\beta, \nu > 0$  and  $Y_* = \{y_*\}$ . This, by Thm. 4.4.3(i), implies  $y^k \rightarrow y_*$ . Therefore we have the following result.

**Theorem 4.6.2.** Suppose that  $f$  and  $g$  satisfy Assumption 4.2, and that (4.D) has a (unique) strong minimizer  $y_*$ . If (4.30) holds in Algorithm 6, then

- (i) the stepsize  $\tau_k = 1$  for all  $k$  sufficiently large,
- (ii) the cost  $\psi(y^k) \rightarrow \inf \psi$   $Q$ -superlinearly,
- (iii) the dual iterates  $y^k \rightarrow y_*$   $Q$ -superlinearly,
- (iv) the primal iterates  $x^k \rightarrow x_*$   $R$ -superlinearly.

*Proof.* We know from Thm.s 4.5.2(ii) and 4.5.4(c) that  $\psi_\gamma$  is twice differentiable with symmetric and positive definite Hessian  $H_* = \nabla^2 \psi_\gamma(y_*)$ . We can expand  $\psi_\gamma$  around  $y_*$  and obtain

$$\begin{aligned} \frac{\psi_\gamma(y^k + d^k) - \inf \psi}{\psi_\gamma(y^k) - \inf \psi} &= \frac{\langle H_*(y^k + d^k - y_*), y^k + d^k - y_* \rangle + o(\|y^k + d^k - y_*\|^2)}{\langle H_*(y^k - y_*), y^k - y_* \rangle + o(\|y^k - y_*\|^2)} \\ &\leq \frac{\|H_*\| \left( \frac{\|y^k + d^k - y_*\|}{\|y^k - y_*\|} \right)^2 + \left( \frac{o(\|y^k + d^k - y_*\|)}{\|y^k - y_*\|} \right)^2}{\lambda_{\min}(H_*) + \left( \frac{o(\|y^k - y_*\|)}{\|y^k - y_*\|} \right)^2} \end{aligned}$$

which vanishes for  $k \rightarrow \infty$ . In particular, eventually  $\psi_\gamma(y^k + d^k) \leq \psi_\gamma(y^k)$  will always hold, proving 4.6.2(i). In turn, since eventually  $\tilde{y}^k = y^k + \tau_k d^k = y^k + d^k$ , using Thm. 4.3.3(ii) and (4.18b) we have

$$\frac{\psi(y^{k+1}) - \inf \psi}{\psi(y^k) - \inf \psi} \leq \frac{\psi_\gamma(\tilde{y}^k) - \inf \psi}{\psi_\gamma(y^k) - \inf \psi} \rightarrow 0,$$

which proves 4.6.2(ii). Moreover, (4.30) reads

$$\|\tilde{y}^k - y_\star\|/\|y^k - y_\star\| \rightarrow 0. \quad (4.31)$$

Now, using nonexpansiveness of  $T_\gamma$  (cf. the proof of [9, Thm. 25.8]) one has

$$\|y^{k+1} - y_\star\| = \|T_\gamma(\tilde{y}^k) - T_\gamma(y_\star)\| \leq \|\tilde{y}^k - y_\star\|$$

which, with (4.31), proves 4.6.2(iii). 4.6.2(iv) follows from 4.6.2(ii) and Lem. 4.A.4.  $\square$

When quasi-Newton directions are computed as in (4.9), superlinear convergence holds provided that the sequence of matrices  $(B_k)_{k \in \mathbb{N}}$  satisfies the Dennis-Moré condition given in the following result. Such condition is satisfied for example by the modified Broyden method (4.10) under standard assumptions of *calm semidifferentiability* of  $R_\gamma$ , see [149, Thm. 6.8].

**Theorem 4.6.3** (Dennis-Moré condition). *Suppose that  $f$  and  $g$  satisfy Assumption 4.2<sup>+</sup>, and that (4.D) has a (unique) strong minimizer  $y_\star$ . If  $(d^k)_{k \in \mathbb{N}}$  is selected according to (4.9), with*

$$\lim_{k \rightarrow \infty} \frac{\|(B_k - JR_\gamma(y_\star))d^k\|}{\|d^k\|} = 0, \quad (4.32)$$

*then  $(d^k)_{k \in \mathbb{N}}$  is superlinearly convergent with respect to  $(y^k)_{k \in \mathbb{N}}$ . In particular, the conclusions of Theorem 4.6.2 hold.*

*Proof.* From Thm.s 4.5.2(i) and 4.5.4(b) we know that  $R_\gamma$  is strictly differentiable, its Jacobian  $J_\star = JR_\gamma(y_\star)$  being nonsingular. Let us denote  $r^k = z^k - Ax^k = R_\gamma(y^k)$  for simplicity. Using (4.9) and (4.32), applying the reverse triangle inequality yields

$$0 \leftarrow \frac{\|r^k - J_\star d^k\|}{\|d^k\|} \geq \frac{\|J_\star B_k^{-1} r^k\|}{\|d^k\|} - \frac{\|r^k\|}{\|d^k\|} \geq \alpha - \frac{\|r^k\|}{\|d^k\|},$$



where  $\alpha = \sqrt{\lambda_{\min}(J_{\star}^{\top} J_{\star})} > 0$  since  $J_{\star}$  is nonsingular. Therefore

$$\liminf_{k \rightarrow \infty} \frac{\|r^k\|}{\|d^k\|} \geq \alpha,$$

and as a consequence  $\|d^k\| \leq (2/\alpha)\|r^k\|$  for all  $k$  sufficiently large. Since  $r^k \rightarrow 0$  by [Thm. 4.4.3\(i\)](#), then  $d^k \rightarrow 0$ . We have

$$0 \leftarrow \frac{r^k - J_{\star}d^k}{\|d^k\|} = \frac{r^k + J_{\star}d^k - R_{\gamma}(y^k + d^k)}{\|d^k\|} + \frac{R_{\gamma}(y^k + d^k)}{\|d^k\|}.$$

The first summand in the above equation tends to zero because of strict differentiability of  $R_{\gamma}$  at  $y_{\star}$ , therefore

$$R_{\gamma}(y^k + d^k)/\|d^k\| \rightarrow 0.$$

By nonsingularity of  $J_{\star}$  then  $\|R_{\gamma}(y)\| \geq \alpha\|y - y_{\star}\|$  for all  $y$  sufficiently close to  $y_{\star}$ , and since  $y^k + d^k \rightarrow y_{\star}$  we have

$$0 \leftarrow \frac{R_{\gamma}(y^k + d^k)}{\|d^k\|} \geq \frac{\alpha\|y^k + d^k - y_{\star}\|}{\|d^k\|} \geq \frac{\alpha\|y^k + d^k - y_{\star}\|}{\|y + d^k - y_{\star}\| + \|y^k - y_{\star}\|}.$$

This implies  $\|y^k + d^k - y_{\star}\|/\|y^k - y_{\star}\| \rightarrow 0$ , i.e.,  $(d^k)_{k \in \mathbb{N}}$  is superlinearly convergent with respect to  $(y^k)_{k \in \mathbb{N}}$ .  $\square$

## 4.7 Simulations

We now present numerical results obtained with the proposed algorithm. Our implementation of NAMA is available online<sup>2</sup>. In all experiments we have considered NAMA with the line-search directions  $(d^k)_{k \in \mathbb{N}}$  computed according to the L-BFGS method, with memory 20, which is able to scale with the problem dimension much better than full quasi-Newton update formulas. All experiments were performed using MATLAB 2016b

---

<sup>2</sup><http://kul-forbes.github.io/ForBES/>

(version 9.1.0) on a MacBook Pro running macOS 10.12, with an Intel Core i5 CPU (2.7 GHz) and 8 GB of memory.

### 4.7.1 Linear MPC

We consider finite horizon, discrete time, linear optimal control problems of the form

$$\underset{\substack{x_0, \dots, x_N \\ u_0, \dots, u_{N-1}}}{\text{minimize}} \sum_{i=0}^{N-1} \ell_i(x_i, u_i) + \ell_N(x_N) \quad (4.33a)$$

$$\text{subject to } x_0 = x_{\text{init}}, \quad (4.33b)$$

$$x_{i+1} = \Phi_i x_i + \Gamma_i u_i + c_i, \quad i = 0, \dots, N-1, \quad (4.33c)$$

where  $x_0, \dots, x_N \in \mathbb{R}^{n_x}$  and  $u_0, \dots, u_{N-1} \in \mathbb{R}^{n_u}$ , and

$$\ell_i(x, u) = q_i(x, u) + g_i(L_i(x, u)), \quad (4.33d)$$

$$\ell_N(x) = q_N(x) + g_N(L_N x). \quad (4.33e)$$

Here the  $q_i$  are strongly convex (typically quadratic), the  $g_i$  are proper, closed, convex functions, while the  $L_i$  are linear mappings, for  $i = 0, \dots, N$ . For example, with a convex set  $C$ , one can set

$$g_i(\cdot) = \delta_C(\cdot) \quad (\text{hard constraints})$$

$$g_i(\cdot) = \alpha \text{dist}_C(\cdot), \quad \alpha > 0, \quad (\text{soft constraints})$$

Set  $C$  here is typically the nonpositive orthant or a box, but can be any other convex set onto which one can efficiently project. When  $C$  is a  $d$ -dimensional box, *i.e.*,  $C = [l_1, u_1] \times \dots \times [l_d, u_d]$ , then one can alternatively model soft constraints as

$$g_i(\cdot) = \sum_{j=1}^d \alpha_i \max \{l_i, \min \{u_i, \cdot\}\}. \quad (4.34)$$

Problem (4.33) takes the form (4.P) by reformulating it as follows (see also [114, 73, 144]). Denote the full sequence of states and inputs as  $\bar{x} =$

$(x_0, u_0, x_1, u_1, \dots, x_N)$ , and let

$$S(p) = \{\bar{x} \mid x_i = \Phi_i x_i + \Gamma_i u_i, x_0 = p\},$$

be the affine subspace of feasible trajectories of the system having initial state  $p$ . Then in (4.P)

$$\begin{aligned} f(\bar{x}) &= \sum_{i=0}^{N-1} q_i(x_i, u_i) + q_N(x_N) + \delta_S(x_{\text{init}})(\bar{x}), \\ g(\bar{z}) &= \sum_{i=0}^N g_i(z_i), \quad A = \text{diag}(L_0, \dots, L_N). \end{aligned}$$

Let us further denote by  $\bar{y} = (y_0, \dots, y_N)$  the dual variable associated with the above problem. In this case, in the alternating minimization step 1 of NAMA, the iterate  $\bar{x}^k$  is obtained by solving

$$\begin{aligned} &\text{minimize } \sum_{i=0}^{N-1} q_i(x_i, u_i) + \langle y_i^k, L_i(x_i, u_i) \rangle \\ &\quad + q_N(x_N) + \langle y_N^k, L_N x_N \rangle. \\ &\text{subject to } x_{i+1} = \Phi_i x_i + \Gamma_i u_i + c_i, \quad i = 0, \dots, N-1. \end{aligned}$$

This is an unconstrained LQR problem whose solution can be efficiently computed with a Riccati-like recursion procedure, in the typical case where  $q_0, \dots, q_N$  are quadratic, see [114, Alg.s 3, 4]. The expensive “factor” step only needs to be performed once, before the main loop of the algorithm takes place. At every iteration one needs to perform merely a forward-backward sweep and no matrix inversions are required. Furthermore

$$\begin{aligned} \bar{z}_i^k &= \text{prox}_{\gamma^{-1}g_i}(\gamma^{-1}y_i^k + L_i(x_i^k, u_i^k)), \quad i = 0, \dots, N-1, \\ \bar{z}_N^k &= \text{prox}_{\gamma^{-1}g_N}(\gamma^{-1}y_N^k + L_N(x_N^k)), \end{aligned}$$

which in the case of hard/soft constraints essentially consist of projec-

tions onto the constrained sets.

## Aircraft control

We applied the proposed method to the AFTI-16 aircraft control problem [13, 73] with  $n_x = 4$  states and  $n_u = 2$  inputs. The dynamics (4.33c), for a sampling time  $T_s = 0.05$  seconds, is given by  $c_i \equiv 0$  and

$$\Phi_i \equiv \begin{bmatrix} 0.9993 & -3.0083 & -0.1131 & -1.6081 \\ 0.0000 & 0.9862 & 0.0478 & 0.0000 \\ 0.0000 & 2.0833 & 1.0089 & 0.0000 \\ 0.0000 & 0.0526 & 0.0498 & 1.0000 \end{bmatrix},$$

$$\Gamma_i \equiv \begin{bmatrix} -0.0804 & -0.6347 \\ -0.0291 & -0.0143 \\ -0.8679 & -0.0917 \\ -0.0216 & -0.0022 \end{bmatrix}.$$

The objective is to drive the *pitch angle* from  $0^\circ$  to  $10^\circ$ , and then back to  $0^\circ$ . We simulated the system for 4 seconds, at the sampling time  $T_s = 0.05$ , using  $N = 50$  and quadratic costs

$$q_i(x, u) = \frac{1}{2} \|x - x_{\text{ref}}\|_Q^2 + \frac{1}{2} \|u\|_R^2, \quad i = 0, \dots, N-1,$$

$$q_N(x) = \frac{1}{2} \|x - x_{\text{ref}}\|_{Q_N}^2,$$

where

$$Q = \text{diag}(10^{-4}, 10^2, 10^{-3}, 10^2),$$

$$Q_N = 100 \cdot Q,$$

$$R = \text{diag}(10^{-2}, 10^{-2}).$$

The reference was set  $x_{\text{ref}} = (0, 0, 0, 10)$  for the first 2 seconds, and  $x_{\text{ref}} = (0, 0, 0, 0)$  for the remaining 2 seconds. Furthermore, we imposed hard box constraints on the inputs, and soft box constraints (4.34) on the states, with weights  $10^6$ , so as to ensure feasibility of the problem. Since soft constraints can be formulated into a QP, by adding linearly penalized nonnegative slack variables, we compared against standard QP solvers.

The dual problem has a condition number of  $10^8$ . To improve the convergence of the algorithms we therefore considered reparametrizing the dual variables according to the *Jacobi scaling*, which consists of a diagonal change of variable (in the dual space) enforcing the (dual) Hessian to have diagonal elements equal to one (see also [126, 73] on the problem of preconditioning fast dual proximal gradient methods). Note that a diagonal change of variable in the dual space simply corresponds to a scaling of the equality constraints, when the problem is equivalently formulated as (P').

The results of the simulations are reported in Table 3. As termination criterion for NAMA and fast AMA we used  $\|R_\gamma(y^k)\|_\infty \leq \epsilon_{\text{tol}} = 10^{-4}$ . We compared NAMA against fast AMA, qpOASES v3.2.0 [62] and the commercial QP solver MOSEK v7.1. We also compared against the cone solvers ECOS v2.0.4 [51], SDPT3 v4.0 [152] and SeDuMi v1.34 [148], all accessed through CVX v2.1 in MATLAB: note that the CPU time for these methods does not include the problem parsing and preprocessing by CVX, but only considers the actual running time of the solvers. Apparently, NAMA greatly improves the convergence performance with respect to fast AMA. When the problem is prescaled, our method outperforms also the other QP and cone solvers considered. One must keep in mind that NAMA was executed using a generic, high-level MATLAB implementation. As computation times become smaller and smaller, overheads due to the runtime environment get more and more relevant in the total CPU time. A tailored, low-level implementation of the same algorithm could significantly decrease the CPU times shown in Table 3: this is also reported in [73], where a speedup of more than a factor 20 is

|          |                          | Iterations |         | CPU time (ms) |         |
|----------|--------------------------|------------|---------|---------------|---------|
|          |                          | avg.       | max.    | avg.          | max.    |
| Fast AMA | (no scaling)             | 12.8 k     | 118.3 k | 5675.3        | 53781.5 |
| NAMA     | (L-BFGS, no scaling)     | 131.8      | 1002    | 138.4         | 974.6   |
| Fast AMA | (Jacobi scaling)         | 203.2      | 922     | 89.8          | 390.3   |
| NAMA     | (L-BFGS, Jacobi scaling) | 14.2       | 57      | 13.7          | 55.5    |
| qpOASES  |                          |            |         | 2925.3        | 4437.8  |
| qpOASES  | (warm-started)           |            |         | 14.3          | 440.2   |
| MOSEK    |                          |            |         | 109.7         | 239.2   |
| ECOS     |                          |            |         | 24.0          | 43.1    |
| SDPT3    |                          |            |         | 645.9         | 1058.3  |
| SeDuMi   |                          |            |         | 122.0         | 190.3   |

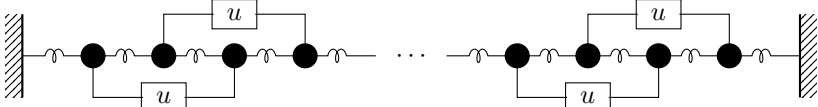
**Table 3:** Aircraft control: performance of the algorithms in the case of the AFTI-16 problem, for  $T_s = 50$  ms and  $N = 50$ . Fast AMA and NAMA were stopped as soon as  $\|R_\gamma(y^k)\|_\infty \leq \epsilon_{\text{tol}} = 10^{-4}$ . Since the problem is ill-conditioned, we also applied the methods by prescaling the dual problem. In NAMA, L-BFGS directions were used with memory 20. NAMA was executed using a generic implementation in MATLAB, while the others QP and cone solvers considered are all implemented in C/C++.

observed using C code generation.

## Oscillating masses

Next, we consider a chain of oscillating masses connected by springs, with both ends attached to walls. The chain is composed of 16 bodies of unit mass, the springs have constant 1 and no damping, and the system is controlled through 8 actuators, each being a force acting on a pair of masses, as depicted in [Figure 14](#). Therefore  $n_x = 32$  (the states are the displacement from the rest position and velocity of each mass) and  $n_u = 8$ . The inputs are constrained in  $[-0.5, +0.5]$ , while the position and velocity of each mass is constrained in  $[-4, +4]$ .

The continuous-time system was discretized with a sampling time  $T_s = 0.5$ . Like in the previous example, we considered quadratic costs with  $Q = Q_N = I_{n_x}$ ,  $R = I_{n_u}$  and hard constraints on state and input.



**Figure 14:** Oscillating masses: schematic representation of the simulated system. The inputs  $u$  are forces acting simultaneously on a pair of masses. In the numerical simulation we considered 16 masses controlled by 8 actuators.

Furthermore, we imposed a quadratic terminal constraint

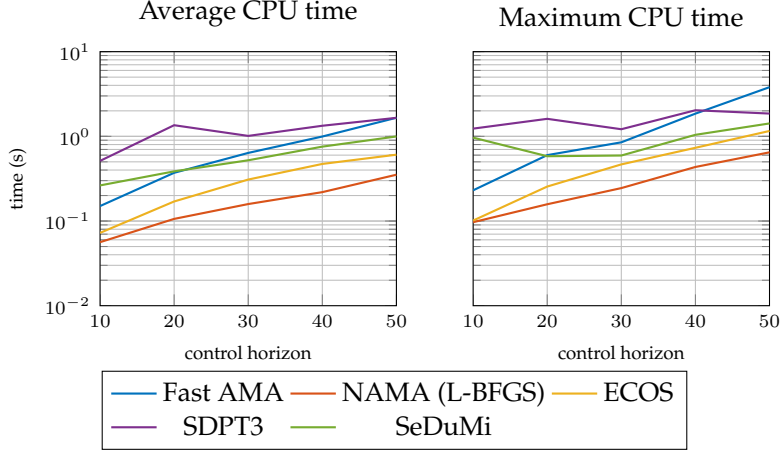
$$\frac{1}{2} \langle Px_N, x_N \rangle \leq \delta, \quad (4.35)$$

where  $P$  solves the Riccati equation related to the discrete-time LQR problem. Constraint (4.35) can be enforced by taking  $L_N$  in (4.33) as the Cholesky factor of  $P$ , so that  $L_N^\top L_N = P$ , and  $g_N$  as the indicator of the Euclidean ball of radius  $\sqrt{\delta}$ . Parameter  $\delta$  is selected so as to ensure that no constraints are violated in such ellipsoidal set.

We simulated different scenarios, each with a different prediction horizon  $N \in \{10, 20, \dots, 50\}$ . For each scenario we selected 50 random initial states  $x_{\text{init}}$  by solving random feasibility problems (*e.g.*, with a cone solver) so as to ensure that a feasible trajectory starting from  $x_{\text{init}}$  exists. Every algorithm was executed with the same set of initial conditions. The results of this experiment are shown in Figure 15. We have compared NAMA against ECOS, SDPT3 and SeDuMi, all accessed through CVX in MATLAB. NAMA outperforms all the other methods considered in this example, and in particular improves considerably over fast AMA, both on average and in the worst case.

## 4.8 Conclusions

In this chapter we presented NAMA, a line-search method for minimizing the sum of two convex functions, one of which is assumed to be



**Figure 15:** Oscillating masses: average and maximum CPU time (in seconds), for increasing prediction horizon and 50 randomly selected initial states. The stopping criterion used for fast AMA and NAMA is  $\|R_\gamma(y^k)\|_\infty \leq \epsilon_{\text{tol}} = 10^{-4}$ .

strongly convex, while the other is composed with a linear transformation. The method is an extension of the classical alternating minimization algorithm (AMA), performing an additional line-search step over the *alternating minimization envelope* associated with the problem. By appropriately selecting the line-search directions, for example according to quasi-Newton methods for solving the optimality conditions  $R_\gamma(y) = 0$ , we have shown that the algorithm converges superlinearly provided that ordinary second order sufficiency conditions hold at the (unique) dual solution. At the same time, the algorithm possesses the same global sublinear and local linear convergence rates as AMA. Numerical experiments with the proposed method on linear MPC problems suggest that NAMA is able to significantly speed up the convergence of AMA, outperforming its accelerated variant and other state-of-the-art solvers even when limited-memory methods, such as L-BFGS, are used to compute the search directions.



## 4.A Additional results

**Lemma 4.A.1.** *Let  $y, w \in \mathbb{R}^m$  and  $\gamma > 0$ . Then,*

$$\begin{aligned} \psi(w) &\geq \psi_\gamma(y) + \frac{\gamma}{2} \|Ax(y) - z_\gamma(y)\|^2 \\ &\quad + \langle z_\gamma(y) - Ax(y), w - y \rangle. \end{aligned} \tag{4.36}$$

*Proof.* By (1.13) we have

$$\begin{aligned} f(x(y)) + f^*(-A^\top w) &\geq -\langle Ax(y), w \rangle, \\ g(z_\gamma(y)) + g^*(w) &\geq \langle z_\gamma(y), w \rangle. \end{aligned}$$

Summing two inequalities and using the definition of  $\psi_\gamma$ , after manipulations one gets the result.  $\square$

**Lemma 4.A.2** (Twice differentiability of  $f^*$ ). *Suppose that  $f$  satisfies Assumption 4.2(i) for the primal-dual solution  $(x_\star, y_\star)$ . Then  $f^*$  is of class  $C^2$  around  $y_\star$ , with*

$$\nabla^2 f^*(y_\star) = \Pi_{S_f} H_f^\dagger \Pi_{S_f}.$$

*Proof.* From [133, Thm. 13.21] we know that  $f^*$  is twice epi-differentiable at  $v$  for  $x \in \partial f^*(v)$  iff  $f$  is twice epi-differentiable at  $x$  for  $v$ , with the relation

$$d^2 f^*(v|x) = [d^2 f(x|v)]^*. \tag{4.37}$$

The cited proof trivially extends to strict twice differentiability, and in fact  $f^*$  turns out to be *strictly* twice epi-differentiable at  $x_\star$ . Since we can assume without loss of generality  $\text{range}(H_f) + S_f^\perp = \mathbb{R}^n$  (see remarks following Assumption 4.2), then applying (4.37) to (4.23) and conjugating  $d^2 f(x_\star | -A^\top y_\star)$  by means of [81, Prop. E.3.2.1] we obtain that function  $f^*$  has purely quadratic second epi-derivative (as opposed to generalized quadratic)

$$d^2 f^*(-A^\top y_\star | x_\star)[w] = \langle \Pi_{S_f} H_f^\dagger \Pi_{S_f} w, w \rangle,$$

which is everywhere finite in particular. The proof concludes invoking [120, Cor. 4.7], which says that  $f^*$  is  $C^2$  around  $-A^\top y_*$ .  $\square$

**Lemma 4.A.3** (Twice epi-differentiability of  $g^*$ ). *Suppose that  $g$  satisfies Assumption 4.2(ii) (respectively, 4.2(ii)<sup>+</sup>) for a primal-dual solution  $(x_*, y_*)$ . Then  $g^*$  is (strictly) twice epi-differentiable at  $y_*$  for  $Ax_*$ , with*

$$d^2 g^*(y_* | Ax_*) = [d^2 g(Ax_* | y_*)]^* = \langle \bar{G} \cdot, \cdot \rangle + \delta_{\bar{S}}(\cdot) \quad (4.38)$$

where  $\bar{G} = \Pi_{S_g} H_g^\dagger \Pi_{S_g}$  and  $\bar{S} = S_g^\perp + \text{range}(H_g)$ .

*Proof.* From [133, Thm. 13.21] we know that  $g^*$  is (strictly) twice epi-differentiable at  $y_*$  for  $Ax_*$ . Similar arguments to those in the proof of Lem. 4.A.2 allow us to conclude that

$$d^2 g^*(y_* | Ax_*)[w] = \langle (\Pi_{S_g} H_g \Pi_{S_g})^\dagger w, w \rangle + \delta_{\bar{S}}(w).$$

Using [14, 6.4.12 i)-ii), 6.1.6 xxxii)] we get  $(\Pi_{S_g} H_g \Pi_{S_g})^\dagger = \Pi_{S_g} H_g^\dagger \Pi_{S_f}$ , which proves (4.38).  $\square$

**Lemma 4.A.4.** *For all  $y \in \mathbb{R}^m$  it holds*

$$\frac{\mu_f}{2} \|x(y) - x_*\|^2 \leq \psi(y) - \inf \psi.$$

*Proof.* From the optimality condition of the problem defining  $x(y)$ , one obtains  $-A^\top y \in \partial f(x(y))$ . Then, by strong convexity of  $f$  one gets

$$f(x(y)) - \langle A^\top y, x_* - x(y) \rangle + \frac{\mu_f}{2} \|x(y) - x_*\|^2 \leq f(x_*).$$

Using (4.13a) in the above inequality we obtain

$$\frac{\mu_f}{2} \|x(y) - x_*\|^2 - \langle Ax_*, y \rangle \leq f(x_*) + f^*(-A^\top y),$$

Using (1.13) on  $g$  we have instead

$$\langle Ax_*, y \rangle \leq g(Ax_*) + g^*(y).$$

Summing the last two inequalities one obtains

$$\frac{\mu_f}{2} \|x(y) - x_*\|^2 \leq f(x_*) + g(Ax_*) + \psi(y),$$

and the claimed bound follows by strong duality.  $\square$

**Lemma 4.A.5.** *Suppose that strict complementarity holds in (4.P), that is,  $0 \in \text{relint } \partial(f + g \circ A)(x_*)$ . Then for any compact set  $U$  there is  $\kappa > 0$  such that*

$$\text{dist}(y, Y_*) \leq \kappa [\text{dist}(-A^\top y, \partial f(x_*)) + \text{dist}(y, \partial g(Ax_*))], \quad \forall y \in U.$$

*Proof.* Consider  $W = \{w \mid -A^\top w \in \partial f(x_*)\}$ . From the optimality conditions (4.2) one has

$$Y_* = W \cap \partial g(Ax_*).$$

If strict complementarity holds, then

$$\begin{aligned} 0 &\in \text{relint} [\partial f(x_*) + A^\top \partial g(Ax_*)] \\ &= \text{relint } \partial f(x_*) + A^\top \text{relint } \partial g(Ax_*). \end{aligned} \tag{4.39}$$

In fact, the first inclusion is due to [Ass. 4.1](#) and [\[132, Thm 23.9\]](#), and the equality is due to [\[132, Thm. 6.6\]](#). If instead  $f$  and  $g$  are piecewise linear-quadratic, then their subdifferentials  $\partial f$  and  $\partial g$  are polyhedral. In both cases, using [\[132, Thm. 6.7\]](#) and [\[8, Cor. 5\]](#) we obtain that the sets  $W$  and  $\partial g(Ax_*)$  are boundedly linearly regular: for any compact  $U$  there is  $\alpha > 0$  such that for all  $y \in U$

$$\text{dist}(y, Y_*) \leq \alpha [\text{dist}(y, W) + \text{dist}(y, \partial g(Ax_*))]. \tag{4.40}$$

Similarly, under the assumed conditions [8, Cor. 5] tells us that the sets

$$L = \mathbb{R}^m \times \text{range}(A^\top), \quad M = \mathbb{R}^m \times \partial f(x_\star),$$

are boundedly linearly regular. Observe that

$$L \cap M = \mathbb{R}^m \times (-A^\top W).$$

Therefore, there is  $\beta > 0$  such that for all  $y \in U$

$$\begin{aligned} \text{dist}(y, W) &\leq \text{dist}((y, -A^\top y), L \cap M) \\ &\leq \beta [\text{dist}((y, -A^\top y), L) + \text{dist}((y, -A^\top y), M)] \\ &\leq \beta \text{dist}(-A^\top y, \partial f(x_\star)), \end{aligned}$$

where the second inequality is due to bounded linear regularity of  $L$  and  $M$ . Using the above inequality in (4.40) yields the result.  $\square$

# Chapter 5

## Fast Douglas-Rachford splitting algorithm

### 5.1 Introduction

In this chapter we consider convex optimization problems of the form

$$\text{minimize } \varphi(x) = f(x) + g(x), \tag{5.1}$$

where  $f : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  and  $g : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  are closed, proper, convex functions with easily computable proximal mappings [129].

A well-known algorithm for solving (5.1) is the Douglas-Rachford splitting (DRS) method [92]. In fact, DRS can be applied to solve the more general problem of finding the zero of two maximal monotone operators. In the special case where the corresponding operators are the

subdifferentials of  $f$  and  $g$ , DRS amounts to the following iterations

$$y^k = \text{prox}_{\gamma f}(x^k), \quad (5.2a)$$

$$z^k = \text{prox}_{\gamma g}(2y^k - x^k), \quad (5.2b)$$

$$x^{k+1} = x^k + \lambda_k(z^k - y^k), \quad (5.2c)$$

where  $\gamma > 0$  and the stepsizes  $\lambda_k \in [0, 2]$  satisfy  $\sum_{k \in \mathbb{N}} \lambda_k(2 - \lambda_k) = +\infty$ . A typical choice for  $\lambda_k$  is to be set equal to 1 for all  $k$ . If the minimum in (5.1) is attained and the relative interiors of the effective domains of  $f$  and  $g$  have a point in common, then it is well known that  $(z^k - y^k)_{k \in \mathbb{N}}$  converges to 0, and  $(x^k)_{k \in \mathbb{N}}$  converges to  $x$  such that  $\text{prox}_{\gamma f}(x) \in \text{argmin } \varphi$  [57, 58, 9]. Therefore  $(y^k)_{k \in \mathbb{N}}$  and  $(z^k)_{k \in \mathbb{N}}$  converge to a solution of (5.1). This general form of DRS was proposed by [57, 58], where it was shown that DRS is a particular case of the proximal point algorithm [129]. Thus DRS converges under very general assumptions. For example, unlike forward-backward splitting (FBS) [37], it does not require differentiability of one of the two summands and parameter  $\gamma$  can take any positive value.

Another well-known application of DRS is for solving problems of the form

$$\text{minimize } f(x) + g(z), \quad (5.3)$$

$$\text{subject to } Ax + Bz = b.$$

Applying DRS to the dual of problem (5.3) leads to the alternating direction method of multipliers (ADMM) [68, 57, 58]. This method has recently received a lot of attention, especially because of its properties with respect to separable objective functions, that make it favorable for large-scale problems and distributed applications [24, 112].

However, when applied to (5.1), the behavior of DRS is quite different compared to standard optimization methods. For example, unlike FBS, DRS is not a descent method, in that the sequence of cost values

$(\varphi(x^k))_{k \in \mathbb{N}}$  may not be monotone decreasing. This is perhaps one of the main reasons why the convergence rate of DRS has not been well understood and convergence rate results were scarce, until very recently. The first convergence result for DRS appeared in [92]. Translated to the setting of solving (5.1), under strong convexity and Lipschitz continuity assumptions for  $f$ , the sequence  $(x^k)_{k \in \mathbb{N}}$  was shown to converge  $Q$ -linearly to the (unique) optimal solution of (5.1). More recently, it was shown that if  $f$  is differentiable then the squared residual  $\|x^k - \text{prox}_{\gamma g}(x^k - \gamma \nabla f(x^k))\|^2$  converges to zero with sublinear rate of  $1/k$  [78]. In [74] convergence rates of order  $1/k$  for the objective values are provided implicitly for DRS under the assumption that both  $f$  and  $g$  have Lipschitz continuous gradients. Under the additional assumption that  $f$  is quadratic, the authors of [74] give an accelerated version with convergence rate  $1/k^2$ . In [46] the authors show global linear convergence for ADMM under a variety of scenarios. Translated in the DRS setting, they require at least  $f$  to be strongly convex with Lipschitz continuous gradient. In [82]  $R$ -linear convergence of the duality gap and primal cost for multiple splitting ADMM under less stringent assumptions is shown, provided that the stepsizes  $\lambda_k$  are sufficiently small. However, the form of the convergence rate is not very informative, since the bound on the stepsizes depends on constants that are very hard to compute. In [70] it is shown that ADMM converges linearly for quadratic programs with the constraint matrix being full rank. However explicit complexity estimates are only provided for the (infrequent) case where the constraint matrix is full row rank. Convergence rates of DRS and ADMM are analyzed under various assumptions in the recent paper [45].

### 5.1.1 Contributions

In this chapter we follow a new approach to the analysis of the convergence properties and complexity estimates of DRS. We show that when  $f$  is twice continuously differentiable, then problem (5.1) is equivalent

to computing a stationary point of a continuously differentiable function, the *Douglas-Rachford Envelope* (DRE). Specifically, DRS is shown to be nothing more than a (scaled) gradient method applied to the DRE. This kind of interpretation is similar to the one offered by the Moreau envelope for the proximal point algorithm and paves the way for deriving new algorithms based on the Douglas-Rachford splitting approach.

A similar idea has been exploited in [113, 117] in order to express another splitting method, the forward-backward splitting, as a gradient method applied to the so-called Forward-Backward Envelope (FBE). There the purpose was use the FBE as a merit function on which to perform Newton-like methods with superlinear local convergence rates to solve non differentiable problems. Here the purpose is instead to analyze the convergence rate properties of Douglas-Rachford splitting by expressing it as a gradient method. Specifically, we show that if  $f$  is convex quadratic (but  $g$  can still be any convex nonsmooth function) then the DRE is convex with Lipschitz continuous gradient, provided that  $\gamma$  is sufficiently small. This covers a wide variety of problems such as quadratic programs,  $\ell_1$  least squares, nuclear norm regularized least squares, image restoration/denoising problems involving total variation minimization norm, etc. This observation makes convergence rate analysis of DRS extremely easy, since it allows us to directly apply the well-known complexity estimates of the gradient method. Furthermore, we discuss the optimal choice of the parameter  $\gamma$  and of the stepsize  $\lambda_k$  defining the method, and devise a method with faster convergence rates by exploiting the acceleration techniques introduced by Nesterov [104],[105, §2.2].

## 5.2 Douglas-Rachford envelope

We indicate by  $X_\star$  the set of optimal solutions to problem (5.1), which we assume to be nonempty. Then  $x_\star \in X_\star$  if and only if [9, Cor. 26.3]



$x_\star = \text{prox}_{\gamma f}(\tilde{x})$ , where  $\tilde{x}$  is a solution of

$$\text{prox}_{\gamma g}(2 \text{prox}_{\gamma f}(x) - x) - \text{prox}_{\gamma f}(x) = 0. \quad (5.4)$$

Let  $\tilde{X}$  be the set of solutions to (5.4). Our goal is to find a continuously differentiable function whose set of stationary points is equal to  $\tilde{X}$ .

Given a function  $h \in \Gamma_0(\mathbb{R}^n)$ , its Moreau envelope  $h^\gamma : \mathbb{R}^n \rightarrow \mathbb{R}$  is differentiable (even if  $h$  is nonsmooth) with  $(1/\gamma)$ -Lipschitz continuous gradient

$$\nabla h^\gamma(x) = \gamma^{-1}(x - \text{prox}_{\gamma h}(x)). \quad (5.5)$$

By using (5.5) we can rewrite (5.4) as

$$\nabla f^\gamma(x) + \nabla g^\gamma(x - 2\gamma \nabla f^\gamma(x)) = 0. \quad (5.6)$$

From now on we make the extra assumption that  $f$  is twice continuously differentiable, with  $L_f$ -Lipschitz continuous gradient. We also assume that  $f$  has strong convexity modulus equal to  $\mu_f \geq 0$ , i.e., function  $f(x) - \frac{\mu_f}{2}\|x\|^2$  is convex. Notice that we allow  $\mu_f$  to be equal to zero, including also the case where  $f$  is not strongly convex. Due to these assumptions we have

$$\|\nabla^2 f(x)\| \leq L_f, \text{ for all } x \in \mathbb{R}^n. \quad (5.7)$$

Moreover, from [87, Prop. 4.1, Th. 4.7] the Jacobian of  $\text{prox}_{\gamma f}$  and the Hessian of  $f^\gamma$  exist everywhere and are related to each other as follows:

$$\nabla \text{prox}_{\gamma f}(x) = (I + \gamma \nabla^2 f(\text{prox}_{\gamma f}(x)))^{-1}, \quad (5.8)$$

$$\nabla^2 f^\gamma(x) = \gamma^{-1}(I - \nabla \text{prox}_{\gamma f}(x)). \quad (5.9)$$

Using (5.7)-(5.9) one can easily show that for any  $d \in \mathbb{R}^n$

$$\frac{\mu_f}{1+\gamma\mu_f} \|d\|^2 \leq \langle d, \nabla^2 f^\gamma(x) d \rangle \leq \frac{L_f}{1+\gamma L_f} \|d\|^2. \quad (5.10)$$

In other words, if  $f$  is twice continuously differentiable with  $L_f$ -Lipschitz

continuous gradient then the eigenvalues of the Hessian of its Moreau envelope are bounded uniformly for every  $x \in \mathbb{R}^n$ .

Next, we premultiply (5.6) by  $(I - 2\gamma\nabla^2 f^\gamma(x))$  to obtain the gradient of what we call the *Douglas-Rachford Envelope (DRE)*:

$$\varphi_\gamma^{\text{DR}}(x) = f^\gamma(x) - \gamma\|\nabla f^\gamma(x)\|^2 + g^\gamma(x - 2\gamma\nabla f^\gamma(x)). \quad (5.11)$$

If  $(I - 2\gamma\nabla^2 f^\gamma(x))$  is nonsingular for every  $x$ , then every stationary point of  $\varphi_\gamma^{\text{DR}}$  is also an element of  $\tilde{X}$ , and vice versa. From (5.10) we obtain

$$\frac{1-\gamma L_f}{1+\gamma L_f}\|d\|^2 \leq \langle d, (I - 2\gamma\nabla^2 f^\gamma(x))d \rangle \leq \frac{1-\gamma\mu_f}{1+\gamma\mu_f}\|d\|^2. \quad (5.12)$$

Therefore whenever  $\gamma < 1/L_f$  or  $\gamma > 1/\mu_f$  (in case where  $\mu_f > 0$ ), finding a stationary point of the DRE (5.11) is equivalent to solving (5.4).

It is convenient now to introduce the following notation:

$$\begin{aligned} P_\gamma(x) &= \text{prox}_{\gamma f}(x), \\ G_\gamma(x) &= \text{prox}_{\gamma g}(2P_\gamma(x) - x), \\ Z_\gamma(x) &= P_\gamma(x) - G_\gamma(x), \end{aligned}$$

so that condition (5.4) is expressed as  $Z_\gamma(x) = 0$ . By (5.9) we can rewrite  $I - 2\gamma\nabla^2 f^\gamma(x) = 2\nabla P_\gamma(x) - I$ , therefore the gradient of the DRE can be expressed as

$$\nabla \varphi_\gamma^{\text{DR}}(x) = \gamma^{-1}(2\nabla P_\gamma(x) - I)Z_\gamma(x). \quad (5.13)$$

The following proposition is instrumental in establishing an equivalence between problem (5.1) and that of minimizing the DRE.

**Proposition 5.2.1.** *The following inequalities hold for any  $\gamma > 0$  and  $x \in \mathbb{R}^n$ :*

$$\varphi_\gamma^{\text{DR}}(x) \leq \varphi(P_\gamma(x)) - \frac{1}{2\gamma}\|Z_\gamma(x)\|^2, \quad (5.14a)$$

$$\varphi_\gamma^{\text{DR}}(x) \geq \varphi(G_\gamma(x)) + \frac{1-\gamma L_f}{2\gamma}\|Z_\gamma(x)\|^2. \quad (5.14b)$$

*Proof.* See Appendix. □

The following fundamental result shows, under the assumption of  $\gamma$  being sufficiently small, that minimizing the DRE, which is real-valued and smooth, is completely equivalent to solving the nonsmooth problem (5.1). Furthermore, the set of stationary points of the DRE, which may not be convex, coincide with the set of its minimizers.

**Theorem 5.2.2.** *If  $\gamma \in (0, 1/L_f)$  then*

$$\begin{aligned} \inf \varphi &= \inf \varphi_\gamma^{\text{DR}}, \\ \operatorname{argmin} \varphi &= P_\gamma(\operatorname{argmin} \varphi_\gamma^{\text{DR}}). \end{aligned}$$

*Proof.* By [9, Cor. 26.3] we know that  $x_\star \in X_\star$  if and only if  $x_\star = P_\gamma(\tilde{x})$ , for some  $\tilde{x} \in \tilde{X}$ , i.e., with  $P_\gamma(\tilde{x}) = G_\gamma(\tilde{x})$ . Putting  $x = \tilde{x}$  in (5.14) one obtains

$$\varphi_\gamma^{\text{DR}}(\tilde{x}) = \varphi(x_\star).$$

When  $\gamma < 1/L_f$ , inequality (5.14b) implies that for all  $x \in \mathbb{R}^n$

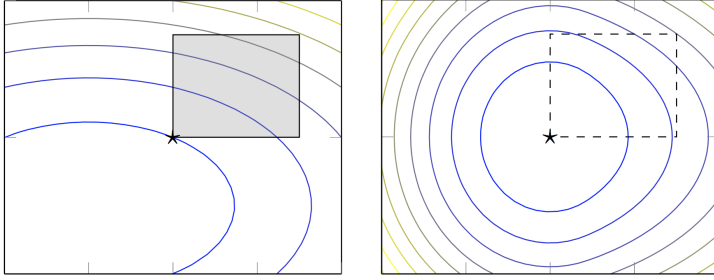
$$\varphi_\gamma^{\text{DR}}(x) \geq \varphi(G_\gamma(x)) \geq \varphi(x_\star) = \varphi_\gamma^{\text{DR}}(\tilde{x}), \quad (5.15)$$

where the last inequality follows from optimality of  $x_\star$ . Therefore the elements of  $\tilde{X}$  are minimizers of  $\varphi_\gamma^{\text{DR}}$  and  $\inf \varphi = \inf \varphi_\gamma^{\text{DR}}$ . They are indeed the only minimizers, for if  $x \notin \tilde{X}$  then  $Z_\gamma(x) \neq 0$  in (5.14b), and the first inequality in (5.15) is strict.  $\square$

### 5.2.1 DRS as a variable-metric gradient method

In simple words, Theorem 5.2.2 tells us that under suitable assumptions on  $\gamma$ , one can employ whichever smooth unconstrained optimization technique for minimizing the DRE and thus solve (5.1). The resulting algorithm will of course bear a close relationship to DRS since the gradient of the DRE, cf. (5.13), is inherently related to a step of DRS, cf. (5.2).

In particular, from the expression (5.13) for  $\nabla \varphi_\gamma^{\text{DR}}$ , one observes that



**Figure 16:** Douglas-Rachford envelope in the case of a box-constrained QP. Left: contour plot of the original quadratic cost, with the feasible region highlighted and the problem solution marked by a star. Right: contour plot of the DRE, whose minimizer coincides with the original problem solution.

Douglas-Rachford splitting can be interpreted as a variable-metric gradient method for minimizing  $\varphi_\gamma^{\text{DR}}$ . Specifically, we have that the  $x$ -iterates defined by (5.2) correspond to

$$x^{k+1} = x^k - \lambda_k D^k \nabla \varphi_\gamma^{\text{DR}}(x^k), \quad (5.16)$$

where

$$D^k = \gamma(2\nabla P_\gamma(x^k) - I)^{-1}. \quad (5.17)$$

We can then exploit all the well-known convergence results of gradient methods to analyze the properties of DRS or propose alternative schemes of it.

## 5.2.2 Connection between DRS and FBS

The DRE reveals an interesting link between Douglas-Rachford splitting and forward-backward splitting, that has remained unnoticed at least to our knowledge. Let us first derive an alternative way of expressing the DRE. Since  $P_\gamma(x) = \operatorname{argmin}_z \{f(z) + \frac{1}{2}\|z - x\|^2\}$  satisfies

$$\nabla f(P_\gamma(x)) + \gamma^{-1}(P_\gamma(x) - x) = 0, \quad (5.18)$$

the gradient of the Moreau envelope of  $f$  becomes

$$\nabla f^\gamma(x) = \gamma^{-1}(x - P_\gamma(x)) = \nabla f(P_\gamma(x)). \quad (5.19)$$

Using (5.18), (5.19) in (5.11) we obtain the following alternative expression for the DRE

$$\varphi_\gamma^{\text{DR}} = f(P_\gamma(x)) - \frac{\gamma}{2} \|\nabla f(P_\gamma(x))\|^2 + g^\gamma(2P_\gamma(x) - x), \quad (5.20)$$

Next, using the definition of  $g^\gamma$  in (5.20), it is possible to express

$$\begin{aligned} \varphi_\gamma^{\text{DR}}(x) = \min_{z \in \mathbb{R}^n} \{ & f(P_\gamma(x)) + \langle \nabla f(P_\gamma(x)), z - P_\gamma(x) \rangle \\ & + g(z) + \frac{1}{2\gamma} \|z - P_\gamma(x)\|^2 \}. \end{aligned} \quad (5.21)$$

Comparing this with the definition of the forward-backward envelope (FBE) introduced in [113]

$$\varphi_\gamma^{\text{FB}}(x) = \min_{z \in \mathbb{R}^n} \left\{ f(x) + \langle \nabla f(x), z - x \rangle + g(z) + \frac{1}{2\gamma} \|z - x\|^2 \right\},$$

it is apparent that the DRE at  $x$  is equal to the FBE evaluated at  $P_\gamma(x)$ :

$$\varphi_\gamma^{\text{DR}}(x) = \varphi_\gamma^{\text{FB}}(P_\gamma(x)).$$

Let us recall here that iterates  $x^{k+1}$  of FBS are obtained by solving the optimization problem appearing in the definition of FBE for  $x = x^k$ . Therefore, it can be easily seen that an iteration of DRS corresponds to a forward-backward step applied to  $\text{prox}_{\gamma f}(x^k)$  (instead of  $x^k$ , as in FBS).

## 5.3 Convergence rate and stepsize selection in DRS

In case  $f$  is convex quadratic, i.e.,  $f(x) = \frac{1}{2} \langle x, Qx \rangle + \langle q, x \rangle$ , with symmetric and positive semidefinite Hessian  $Q \in \mathbb{R}^{n \times n}$  and  $q \in \mathbb{R}^n$ , we have

$$P_\gamma(x) = (I + \gamma Q)^{-1}(x - \gamma q), \quad (5.22)$$

$$\nabla P_\gamma(x) = (I + \gamma Q)^{-1}. \quad (5.23)$$

We now have  $\mu_f = \lambda_{\min}(Q)$  and  $L_f = \lambda_{\max}(Q)$ . It turns out that in this case, under the already mentioned assumption  $\gamma < 1/L_f$ , the DRE is convex.

**Theorem 5.3.1.** *Suppose that  $f$  is convex quadratic. If  $\gamma < 1/L_f$ , then  $\varphi_\gamma^{\text{DR}}$  is convex with  $L_{\varphi_\gamma^{\text{DR}}}$ -Lipschitz continuous gradient and convexity modulus  $\mu_{\varphi_\gamma^{\text{DR}}}$  given by*

$$L_{\varphi_\gamma^{\text{DR}}} = \frac{1 - \gamma\mu_f}{1 + \gamma\mu_f} \gamma^{-1}, \quad (5.24)$$

$$\mu_{\varphi_\gamma^{\text{DR}}} = \min \left\{ \frac{(1 - \gamma\mu_f)\mu_f}{(1 + \gamma\mu_f)^2}, \frac{(1 - \gamma L_f)L_f}{(1 + \gamma L_f)^2} \right\}. \quad (5.25)$$

*Proof.* Using (5.13), (5.23), (5.12) and Lemma 5.A.2 in the Appendix, we obtain

$$\begin{aligned} \|\nabla \varphi_\gamma^{\text{DR}}(x_1) - \nabla \varphi_\gamma^{\text{DR}}(x_2)\| &\leq \gamma^{-1} \|2(I + \gamma Q)^{-1} - I\| \cdot \|Z_\gamma(x_1) - Z_\gamma(x_2)\| \\ &\leq \left( \frac{2}{1 + \gamma\mu_f} - 1 \right) \gamma^{-1} \|x_1 - x_2\|. \end{aligned}$$

Next, due to the form of  $P_\gamma$ , cf. (5.22) it is evident that  $f(P_\gamma(x)) - \frac{\gamma}{2} \|\nabla f(P_\gamma(x))\|^2$  is quadratic with Hessian

$$H = (I + \gamma Q)^{-1}(I - \gamma Q)Q(I + \gamma Q)^{-1}.$$

The eigenvalues of  $H$  are given by  $\frac{(1 - \gamma\lambda_i)\lambda_i}{(1 + \gamma\lambda_i)^2}$ , where  $\lambda_i, i = 1, \dots, n$  are

the eigenvalues of  $Q$ . Consider the function

$$\zeta(\lambda) = \frac{(1-\gamma\lambda)\lambda}{(1+\gamma\lambda)^2}.$$

If  $\gamma < 1/L_f$ ,  $\zeta$  is concave and its minimum is attained in one of the two endpoints of the interval  $[\mu_f, L_f]$ . The minimum eigenvalue of  $f(P_\gamma(x)) - \frac{\gamma}{2}\|\nabla f(P_\gamma(x))\|^2$  is then given by (5.25). On the other hand, function  $g^\gamma(\cdot - 2\gamma\nabla f^\gamma(\cdot))$  is convex as the composition of the convex function  $g^\gamma$  with an affine map. Therefore, the DRE as expressed by (5.20), is the sum of two functions, one of them being (strongly) convex with modulus  $\mu_{\varphi_\gamma^{\text{DR}}}$  and the other convex. Hence it is (strongly) convex with modulus  $\mu_{\varphi_\gamma^{\text{DR}}}$ .  $\square$

Therefore, under the assumptions of Theorem 5.3.1, we can exploit the well-known results on the convergence of the gradient method for convex problems. To do so, note that when  $f$  is quadratic,  $P_\gamma$  is linear and the scaling matrix  $D^k$  defined in (5.17) is constant, *i.e.*,

$$D^k \equiv D = \gamma[2(I + \gamma Q)^{-1} - I]^{-1}.$$

Consider the linear change of variables  $x = Sw$ , where  $S = D^{1/2}$ . Note that

$$\lambda_{\min}(D) = \gamma \frac{1 + \gamma\mu_f}{1 - \gamma\mu_f}, \quad \lambda_{\max}(D) = \gamma \frac{1 + \gamma L_f}{1 - \gamma L_f}, \quad (5.26)$$

so if  $\gamma < 1/L_f \leq 1/\mu_f$  then matrix  $D$  is positive definite and  $S$  is well defined.

In the new variable  $w$ , the scaled gradient iterations (5.16) correspond to the (unscaled) gradient method applied to the preconditioned problem

$$\text{minimize } h(w) = \varphi_\gamma^{\text{DR}}(Sw).$$

Indeed, the gradient method applied on  $h$  is

$$w^{k+1} = w^k - \lambda_k \nabla h(w^k) \quad (5.27)$$

Multiplying by  $S$  and using  $\nabla h(w^k) = S\nabla\varphi_\gamma^{\text{DR}}(Sw^k)$ , we obtain

$$x^{k+1} = x^k - \lambda_k D \nabla \varphi_\gamma^{\text{DR}}(x^k).$$

Recalling (5.13), this becomes  $x^{k+1} = x^k - \lambda_k Z_\gamma(x^k)$ , which is exactly DRS, cf. (5.2). From now on we will indicate by  $\tilde{w}$  a minimizer of  $h$ , so that  $\tilde{w} = S\tilde{x}$  for some  $\tilde{x} \in \tilde{X}$ . From Theorem 5.3.1 we know that if  $\gamma < 1/L_f$  then  $\varphi_\gamma^{\text{DR}}$  is convex with Lipschitz continuous gradient, and so is  $h$ . In particular,

$$\mu_h = \lambda_{\min}(D)\mu_{\varphi_\gamma^{\text{DR}}}, \quad (5.28)$$

$$L_h = \lambda_{\max}(D)L_{\varphi_\gamma^{\text{DR}}} = \frac{1 + \gamma L_f}{1 - \gamma L_f}. \quad (5.29)$$

**Theorem 5.3.2.** *For convex quadratic  $f$ , if  $\gamma < 1/L_f$  and*

$$\lambda_k = \lambda = (1 - \gamma L_f)/(1 + \gamma L_f) \quad (5.30)$$

*then the sequence of iterates generated by (5.2a)-(5.2c) satisfies*

$$\varphi(z^{k+1}) - \inf \varphi \leq \frac{1}{(2\gamma\lambda)k} \|x^0 - \tilde{x}\|^2.$$

*Proof.* Douglas-Rachford splitting (5.2) corresponds to the gradient descent iterations (5.27). So by setting  $\lambda = 1/L_h$  one has:

$$h(w^k) - h(\tilde{w}) \leq \frac{L_h}{2k} \|w^0 - \tilde{w}\|^2,$$

see for example [17, Prop. 6.10.2]. Applying the substitution  $x = Sw$ , and considering that

$$\lambda_{\max}^{-1}(D)\|x\|^2 \leq \|x\|_{D^{-1}}^2 \leq \lambda_{\min}^{-1}(D)\|x\|^2, \quad \forall x \in \mathbb{R}^n \quad (5.31)$$



one obtains

$$\begin{aligned}
\varphi_\gamma^{\text{DR}}(x^k) - \varphi_\gamma^{\text{DR}}(\tilde{x}) &\leq \frac{L_h}{2k} \|x^0 - \tilde{x}\|_{D^{-1}}^2 \\
&\leq \frac{1}{2k} \frac{1 + \gamma L_f}{(1 - \gamma L_f)} \frac{1}{\lambda_{\min}(D)} \|x^0 - \tilde{x}\|^2 \\
&= \frac{1}{2k} \frac{1 + \gamma L_f}{\gamma(1 - \gamma L_f)} \|x^0 - \tilde{x}\|^2,
\end{aligned}$$

where the last equality holds considering (5.26). The claim follows by  $z^k = G_\gamma(x^k)$ , Theorem 5.2.2 and inequality (5.14b).  $\square$

From Theorem 5.3.2 we easily obtain the following optimal  $\gamma$ :

$$\gamma_\star = \operatorname{argmin}_\gamma \frac{1 + \gamma L_f}{\gamma(1 - \gamma L_f)} = \frac{\sqrt{2} - 1}{L_f}. \quad (5.32)$$

For this particular value of  $\gamma_\star$  the stepsize becomes  $\lambda_k = \sqrt{2} - 1$ . In the strongly convex case we instead obtain the following stronger result.

**Theorem 5.3.3.** *If  $\mu_f > 0$  and  $\lambda_k = \lambda \in (0, 2/(L_h + \mu_h)]$  then*

$$\|y^k - x_\star\|^2 \leq \frac{\lambda_{\max}(D)}{\lambda_{\min}(D)} \left(1 - \frac{2\lambda\mu_h L_h}{\mu_h + L_h}\right)^k \|x^0 - \tilde{x}\|^2.$$

*Proof.* Just like in the proof of Theorem 5.3.2, iteration (5.27) is the standard gradient method applied to  $h$ . If  $f$  is strongly convex then we have, using (5.25) and (5.28), that also  $h$  is strongly convex. From [105, Th. 2.1.15] we have

$$\|w^k - \tilde{w}\|^2 \leq \left(1 - \frac{2\lambda\mu_h L_h}{\mu_h + L_h}\right)^k \|w^0 - \tilde{w}\|^2.$$

Applying the substitution  $x = Sw$  we get

$$\|x^k - \tilde{x}\|_{D^{-1}}^2 \leq \left(1 - \frac{2\lambda\mu_h L_h}{\mu_h + L_h}\right)^k \|x^0 - \tilde{x}\|_{D^{-1}}^2.$$

The thesis follows considering (5.31) and that

$$\|y^k - x_\star\|^2 = \|\text{prox}_{\gamma f}(x^k) - \text{prox}_{\gamma f}(\tilde{x})\|^2 \leq \|x^k - \tilde{x}\|^2,$$

where the equality holds since  $x_\star = \text{prox}_{\gamma f}(\tilde{x})$ , and the inequality by nonexpansiveness of  $\text{prox}_{\gamma f}$ .  $\square$

## 5.4 Fast Douglas-Rachford splitting

We have shown that DRS is equivalent to the gradient method minimizing  $h(w) = \varphi_\gamma^{\text{DR}}(Sw)$ . In the quadratic case, since for  $\gamma < 1/L_f$  we know that  $\varphi_\gamma^{\text{DR}}(x)$  is convex, we can as well apply the optimal first order methods due to Nesterov [104], [105, §2.2] to the same problem. This way we obtain a *fast Douglas-Rachford splitting* method. The scheme is as follows: given  $u^0 = x^0 \in \mathbb{R}^n$ , iterate

$$\begin{aligned} y^k &= \text{prox}_{\gamma f}(u^k), \\ z^k &= \text{prox}_{\gamma g}(2y^k - u^k), \\ x^{k+1} &= u^k + \lambda_k(z^k - y^k), \\ u^{k+1} &= x^{k+1} + \beta_k(x^{k+1} - x^k). \end{aligned} \tag{5.33}$$

We have the following estimates regarding the convergence rate of iterations (5.33), whose proofs are based on [105].

**Theorem 5.4.1.** *For convex quadratic  $f$ , if  $\gamma < 1/L_f$ ,  $\lambda_k$  are given by (5.30) and*

$$\beta_k = \begin{cases} 0 & \text{if } k = 0, \\ \frac{k-1}{k+2} & \text{if } k > 0, \end{cases}$$

*then the sequence of iterates generated by (5.33) satisfies*

$$\varphi(z^k) - \inf \varphi \leq \frac{2}{\gamma \lambda(k+2)^2} \|x^0 - \tilde{x}\|^2.$$

*Proof.* The iterations correspond to the optimal method described in [17, §6.10.2], applied to  $h$ . By [17, Prop. 6.10.3] the iterates satisfy

$$h(w^k) - h(\tilde{w}) \leq \frac{2L_h}{(k+2)^2} \|w^0 - \tilde{w}\|^2.$$

Switching to the variable  $x = Sw$  we get

$$\begin{aligned} \varphi_\gamma^{\text{DR}}(x^k) - \varphi_\gamma^{\text{DR}}(\tilde{x}) &\leq \frac{2L_h}{(k+2)^2} \|x^0 - \tilde{x}\|_{D^{-1}}^2 \\ &\leq \frac{1}{\lambda_{\min}(D)} \frac{2L_h}{(k+2)^2} \|x^0 - \tilde{x}\|^2 \\ &= \frac{\lambda_{\max}(D)}{\lambda_{\min}(D)} \frac{2L_{\varphi_\gamma^{\text{DR}}}}{(k+2)^2} \|x^0 - \tilde{x}\|^2 \\ &= \frac{1 + \gamma L_f}{\gamma(1 - \gamma L_f)} \frac{2}{(k+2)^2} \|x^0 - \tilde{x}\|^2. \end{aligned}$$

Since  $z^k = G_\gamma(x^k)$ , the result follows by invoking inequality (5.14b) and Theorem 5.2.2.  $\square$

The optimal choice for  $\gamma$  is again  $\gamma_\star = (\sqrt{2} - 1)/L_f$ . We similarly obtain complexity bounds for the strongly convex case, as described in the following result.

**Theorem 5.4.2.** *If  $f$  is strongly convex quadratic,  $\gamma < 1/L_f$ ,  $\lambda_k$  are given by (5.30) and*

$$\beta_k = \frac{1 - \sqrt{\mu_h/L_h}}{1 + \sqrt{\mu_h/L_h}},$$

*then the sequence of iterates generated by (5.33) satisfies*

$$\varphi(z^k) - \inf \varphi \leq \frac{L_h}{\lambda_{\min}(D)} \left(1 - \sqrt{\frac{\mu_h}{L_h}}\right)^k \|x^0 - x_\star\|^2.$$

*Proof.* The proof proceeds similarly to the previous one. The algorithm corresponds to iterations [105, Eq. 2.2.9] applied to  $h$ , and [105, Th. 2.2.3]

tells us that

$$h(w^k) - h(\tilde{w}) \leq L_h \left(1 - \sqrt{\frac{\mu_h}{L_h}}\right)^k \|w^0 - \tilde{w}\|^2.$$

The latter is equivalent to

$$\begin{aligned} \varphi_\gamma^{\text{DR}}(x^k) - \varphi_\gamma^{\text{DR}}(\tilde{x}) &\leq L_h \left(1 - \sqrt{\frac{\mu_h}{L_h}}\right)^k \|x^0 - \tilde{x}\|_{D^{-1}}^2 \\ &\leq \frac{L_h}{\lambda_{\min}(D)} \left(1 - \sqrt{\frac{\mu_h}{L_h}}\right)^k \|x^0 - \tilde{x}\|^2. \end{aligned}$$

Again,  $z^k = G_\gamma(x^k)$ , Theorem 5.2.2 and inequality (5.14b) complete the result.  $\square$

## 5.5 Simulations

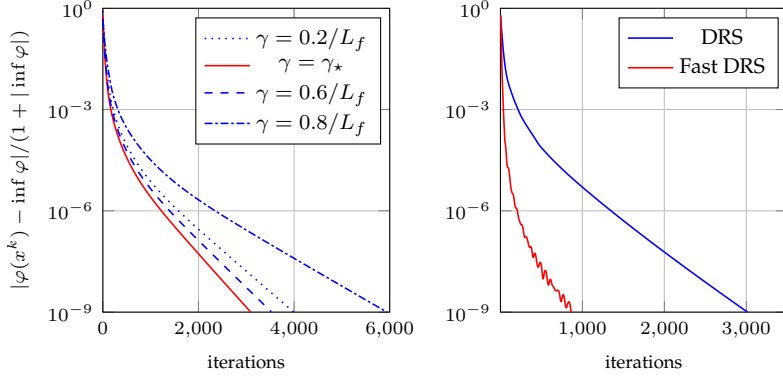
### 5.5.1 Box-constrained QP

We tested our analysis against numerical results obtained by applying the considered methods to the following box-constrained convex quadratic program

$$\begin{aligned} &\text{minimize } \frac{1}{2} \langle x, Qx \rangle + \langle q, x \rangle \\ &\text{subject to } l \leq x \leq u, \end{aligned}$$

where  $Q \in \mathbb{R}^{n \times n}$  is symmetric and positive semidefinite, while  $q, l, u \in \mathbb{R}^n$ . The problem is expressed in composite form by setting

$$f(x) = \frac{1}{2} \langle x, Qx \rangle + \langle q, x \rangle, \quad g(x) = \delta_{[l, u]}(x),$$



**Figure 17:** Box QP. Left: DRS applied to a randomly generated problem with  $n = 500$ , for different values of  $\gamma$ . Right: comparison between DRS and its accelerated variant, for  $\gamma = \gamma_*$

where  $\delta_C$  is the indicator function of the convex set  $C$ . As it was pointed out in Section 5.3, the proximal mapping associated with  $f$  is linear

$$\text{prox}_{\gamma f}(x) = (I + \gamma Q)^{-1}(x - \gamma q).$$

The proximal mapping associated with  $g$  is simply the projection onto the  $[l, u]$  box,  $\text{prox}_{\gamma g}(x) = \Pi_{[l, u](x)}$ . Tests were performed on problems generated randomly as described in [75]. In Figure 17 we illustrate the performance of DRS for different choices of the parameter  $\gamma$ , and compares the standard DRS to the accelerated method (5.33)

### 5.5.2 Lasso

The well-known  $\ell_1$ -regularized least squares problem consists of finding a sparse solution to an underdetermined linear system. The goal is achieved by solving

$$\text{minimize } \frac{1}{2} \|Ax - b\|_2^2 + \rho \|x\|_1,$$

where  $A \in \mathbb{R}^{m \times n}$  and  $b \in \mathbb{R}^m$ . The regularization parameter  $\rho$  modulates between a low residual  $\|Ax - b\|_2^2$  and a sparse solution. In this case the proximal mapping with respect to  $f$  is

$$\text{prox}_{\gamma f}(x) = (A^\top A + \gamma^{-1}I)^{-1}(A^\top b + \gamma^{-1}x),$$

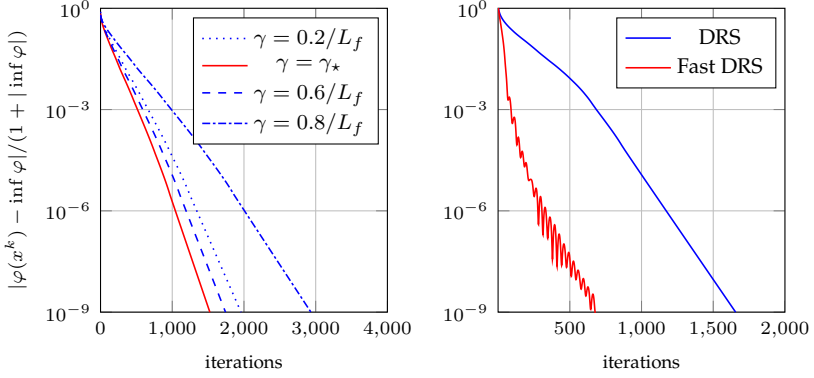
while  $\text{prox}_{\gamma g}$  is the following *soft-thresholding* operator,

$$[\text{prox}_{\gamma g}(x)]_i = \text{sign}(x_i) \cdot \max\{0, |x_i| - \gamma\rho\}, \quad i = 1, \dots, n.$$

Random problems were generated according to [97], and the results are shown in Figure 18, where we compare different choices for  $\gamma$  and the fast Douglas-Rachford iterations. For this type of problems one is often interested in obtaining the solution for a sequence of values of the regularization parameter  $\lambda$ , computing the so-called *regularization path*. In this case, both DRS and fast DRS can be warm-started by directly using the solution to one problem as initial iterate of the next one: this has been shown to accelerate the convergence of other algorithms, see for example [142]. A study of how the proposed fast DRS benefits from warm-starting is subject of future investigation.

## 5.6 Conclusions and future work

In this chapter we dealt with convex composite minimization problems. We introduced a continuously differentiable function, the Douglas-Rachford Envelope (DRE). Its minimizers, under suitable assumptions, are in a one-to-one correspondence with the solutions of the original convex composite optimization problem. We observed how the DRS iterations, for finding zeros of the sum of two maximal monotone operators  $F$  and  $G$ , are equivalent to a scaled unconstrained gradient method applied to the DRE, when  $F = \partial f$  and  $G = \partial g$  and  $f$  is twice continuously differentiable with Lipschitz continuous gradient. This allowed us to apply



**Figure 18:** Lasso. Left: comparison of different choices of  $\gamma$  for a random problem with  $m = 100, n = 1000$ . Right: DRS and its accelerated variant, with  $\gamma = \gamma_*$ .

well-known results of smooth unconstrained optimization to analyze the convergence of DRS in the particular case of  $f$  being convex quadratic. Moreover, we have been able to apply and analyze optimal first-order methods and obtain a fast Douglas-Rachford splitting method. Ongoing work on this topic include exploiting the illustrated results to study convergence properties of ADMM.

## 5.A Proofs and additional results

We provide here all the proofs and technical lemmas omitted in the chapter. The following lemma is instrumental in proving [Proposition 5.2.1](#).

**Lemma 5.A.1.** *Suppose that  $h : \mathbb{R}^n \rightarrow \overline{\mathbb{R}}$  is proper, closed, convex. Then for all  $y \in \mathbb{R}^n, z \in \mathbb{R}^n$*

$$\begin{aligned} h(z) + \frac{1}{2\gamma} \|z - y\|^2 &\geq h(\text{prox}_{\gamma h}(y)) + \frac{1}{2\gamma} \|\text{prox}_{\gamma h}(y) - y\|^2 \\ &\quad + \frac{1}{2\gamma} \|z - \text{prox}_{\gamma h}(y)\|^2. \end{aligned}$$

*Proof.* Function  $\phi(z) = \frac{1}{2\gamma}\|z - y\|^2$  is strongly convex with modulus  $\gamma^{-1}$ . For any  $v \in \partial h(y_\gamma)$  we have, by strong convexity of  $h(z) + \phi(z)$ ,

$$\begin{aligned} h(z) + \phi(z) &= h(z) + \frac{1}{2\gamma}\|z - y\|^2 \\ &\geq h(\text{prox}_{\gamma h}(y)) + \frac{1}{2\gamma}\|\text{prox}_{\gamma h}(y) - y\|^2 + \frac{1}{2\gamma}\|z - \text{prox}_{\gamma h}(y)\|^2 \\ &\quad + \left\langle v + \frac{1}{\gamma}(\text{prox}_{\gamma h}(y) - y), z - \text{prox}_{\gamma h}(y) \right\rangle. \end{aligned}$$

The result follows by considering  $v = \gamma^{-1}(y - \text{prox}_{\gamma h}(y))$ , which is an element of  $\partial h(y_\gamma)$  by the optimality condition for  $\text{prox}_{\gamma h}(y)$ .  $\square$

*Proof of Proposition 5.2.1.* Due to (5.21), an alternative expression for the DRE is the following

$$\begin{aligned} \varphi_\gamma^{\text{DR}}(x) &= f(P_\gamma(x)) + g(G_\gamma(x)) + \frac{1}{2\gamma}\|G_\gamma(x) - P_\gamma(x)\|^2 \\ &\quad + \gamma^{-1} \langle G_\gamma(x) - P_\gamma(x), x - P_\gamma(x) \rangle. \end{aligned} \quad (5.34)$$

In order to obtain (5.14a), apply Lemma 5.A.1 for  $h = g, y = 2P_\gamma(x) - x$ . We have that for all  $z \in \mathbb{R}^n$

$$\begin{aligned} g(z) + \frac{1}{2\gamma}\|z - (2P_\gamma(x) - x)\|^2 &\geq g(G_\gamma(x)) + \frac{1}{2\gamma}\|G_\gamma(x) - (2P_\gamma(x) - x)\|^2 \\ &\quad + \frac{1}{2\gamma}\|z - G_\gamma(x)\|^2. \end{aligned}$$

Putting  $z = P_\gamma(x)$  in the above,

$$\begin{aligned} g(P_\gamma(x)) + \frac{1}{2\gamma}\|x - P_\gamma(x)\|^2 &\geq g(G_\gamma(x)) + \frac{1}{2\gamma}\|P_\gamma(x) - G_\gamma(x)\|^2 \\ &\quad + \frac{1}{2\gamma}\|G_\gamma(x) - P_\gamma(x) + x - P_\gamma(x)\|^2 \\ &= g(G_\gamma(x)) + \frac{1}{2\gamma}\|G_\gamma(x) - P_\gamma(x)\|^2 \\ &\quad + \frac{1}{2\gamma}\|x - P_\gamma(x)\|^2 + \frac{1}{2\gamma}\|P_\gamma(x) - G_\gamma(x)\|^2 \\ &\quad + \gamma^{-1} \langle G_\gamma(x) - P_\gamma(x), x - P_\gamma(x) \rangle. \end{aligned}$$



Therefore,

$$\begin{aligned} g(P_\gamma(x)) &\geq g(G_\gamma(x)) + \frac{1}{2\gamma}\|G_\gamma(x) - P_\gamma(x)\|^2 + \frac{1}{2\gamma}\|P_\gamma(x) - G_\gamma(x)\|^2 \\ &\quad + \gamma^{-1} \langle G_\gamma(x) - P_\gamma(x), x - P_\gamma(x) \rangle. \end{aligned}$$

Adding  $f(P_\gamma(x))$  to both sides,

$$\begin{aligned} \varphi(P_\gamma(x)) &\geq f(P_\gamma(x)) + g(G_\gamma(x)) + \frac{1}{2\gamma}\|G_\gamma(x) - P_\gamma(x)\|^2 \\ &\quad + \frac{1}{2\gamma}\|P_\gamma(x) - G_\gamma(x)\|^2 + \gamma^{-1} \langle G_\gamma(x) - P_\gamma(x), x - P_\gamma(x) \rangle. \end{aligned}$$

We obtain (5.14a) by recalling (5.34). Inequality (5.14b) is obtained as follows,

$$\begin{aligned} \varphi(G_\gamma(x)) &= f(G_\gamma(x)) + g(G_\gamma(x)) \\ &\leq f(P_\gamma(x)) + g(G_\gamma(x)) + \langle \nabla f(P_\gamma(x)), G_\gamma(x) - P_\gamma(x) \rangle \\ &\quad + \frac{L_f}{2}\|G_\gamma(x) - P_\gamma(x)\|^2 \\ &= f(P_\gamma(x)) + g(G_\gamma(x)) + \gamma^{-1} \langle G_\gamma(x) - P_\gamma(x), x - P_\gamma(x) \rangle \\ &\quad + \frac{L_f}{2}\|G_\gamma(x) - P_\gamma(x)\|^2 \\ &= \varphi_\gamma^{\text{DR}}(x) - \frac{1-\gamma L_f}{2\gamma}\|G_\gamma(x) - P_\gamma(x)\|^2, \end{aligned}$$

where the first inequality follows from the Lipschitz continuity of  $\nabla f$  and the last equality from (5.34).  $\square$

The next basic result is used in the proof of Theorem 5.3.1.

**Lemma 5.A.2.** *Mapping  $Z_\gamma : \mathbb{R}^n \rightarrow \mathbb{R}^n$  is nonexpansive.*

*Proof.* We can express  $Z_\gamma$  as

$$Z_\gamma(x) = \frac{1}{2}(x - T(x)),$$

where  $T = R_{\gamma\partial g} \circ R_{\gamma\partial f}$  and  $R_{\gamma\partial f}, R_{\gamma\partial g}$  are called *reflected resolvent* [9, §23] of  $\partial f$  and  $\partial g$ , respectively. Reflected resolvents of maximal monotone mappings (such as the subdifferential of a convex function) are known

to be nonexpansive [9, Cor. 23.10], and so is their composition  $T$ . Then we have

$$\|T(x_1) - T(x_2)\| \leq \|x_1 - x_2\|,$$

for all  $x_1, x_2 \in \mathbb{R}^n$ , or

$$\| -2(Z_\gamma(x_1) - Z_\gamma(x_2)) + (x_1 - x_2) \| \leq \|x_1 - x_2\|.$$

Using the reverse triangle inequality

$$2\|Z_\gamma(x_1) - Z_\gamma(x_2)\| - \|x_1 - x_2\| \leq \|x_1 - x_2\|,$$

or

$$\|Z_\gamma(x_1) - Z_\gamma(x_2)\| \leq \|x_1 - x_2\|,$$

*i.e.*,  $Z_\gamma$  is nonexpansive. □

## Chapter 6

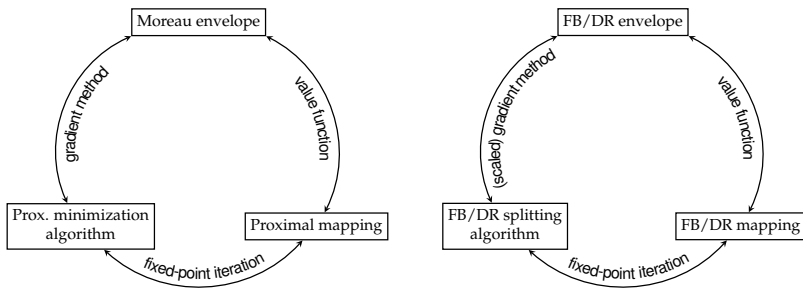
# Conclusions and outlook

In this thesis we introduced *proximal envelopes*, powerful tools for nonsmooth, nonconvex optimization, that

- (i) allow to reformulate structured, nonsmooth optimization problems as smooth, unconstrained ones;
- (ii) give an interpretation of proximal splitting methods as a (scaled) gradient methods;
- (iii) can be evaluated with the same computational effort as one iteration of the correspondent proximal algorithm.

Because of these facts, proximal envelopes generalize and extend the well-known concept of Moreau envelope to structured problems: this analogy is schematically depicted in [Figure 19](#).

In [Chapter 2](#) we have introduced and analyzed the *forward-backward envelope* (FBE). Through the analysis of (twice) differentiability of the FBE (which can be guaranteed by assuming mild, generalized second-order properties on the nonsmooth cost function), we were able to obtain second-order sufficiency conditions which guarantee superlinear convergence of Newton-type methods. Then we proposed MinFBE, an al-



**Figure 19:** The well known relationship between Moreau envelope, proximal mapping and proximal minimization algorithm (left) is completely analogous to that relating the FBE/DRE to the corresponding splitting algorithms (right).

algorithmic scheme for nonconvex problems that performs a line-search over the FBE to obtain fast asymptotic convergence. The algorithm is a descent method, and as such requires gradient evaluations of the FBE to compute descent directions. We were able to show global convergence under the Kurdyka-Łojasiewicz assumption, and a global sublinear rate in the convex case. Furthermore, we showed superlinear convergence when the search directions satisfy the Dennis-Moré condition, as it is the case for quasi-Newton formulas.

In Chapter 3 we proposed a simpler algorithmic scheme, PANOC. This merely requires evaluations of the fixed-point residual: in particular, no gradient information on the FBE is needed, since no descent direction is required in the line-search. Nevertheless, fast convergence directions can be computed in this case by looking at the problem as a nonlinear, nonsmooth system of equations given by the first-order necessary conditions (rather than a minimization problem). This approach is particularly appealing in cases where evaluating gradients of the FBE can be computationally expensive, such as in nonlinear MPC problems.

In Chapter 4 we considered convex, equality constrained problems. We observed that the FBE is closely related to the augmented Lagrangian

| Algorithm                      | MinFBE   | PANOC | NAMA     |
|--------------------------------|----------|-------|----------|
| Requires $f \in C^2$           | yes      | no    | no       |
| FB steps per iter. (best case) | 2        | 1     | 2        |
| Global rate (convex problems)  | $O(1/k)$ | -     | $O(1/k)$ |
| Superlinear convergence        | yes      | yes   | yes      |

**Table 4:** Comparison between the proposed forward-backward algorithmic schemes. PANOC and NAMA are conceptually simpler than MinFBE, in that they do not require second order information on the smooth term. Furthermore PANOC only performs one forward-backward operation per iteration in the best case; MinFBE and NAMA, on the other hand, converge with a sublinear global convergence rate (in the convex case) thanks to the additional forward-backward step per iteration.

function and the alternating minimization algorithm (AMA) for convex problems with equality constraints: in fact, the two functions are dual to each other, strengthening the analogy with the Moreau even further. The first- and second-order analysis of Chapter 2 was here extended to the dual case. Finally, an algorithm was proposed which extends AMA by performing Newton-type dual updates: NAMA. We proved sublinear and linear convergence for the algorithm, under mild assumptions, and superlinear convergence for directions satisfying the Dennis-Moré conditions.

Table 4 summarizes the main features distinguishing the algorithmic schemes of Chapters 2 to 4. These are implemented in a generic MATLAB software package, ForBES, that easily allows operating with FBE to define algorithms. The package is available online.<sup>1</sup>

In Chapter 5 the *Douglas-Rachford splitting* (DRS) was taken into account, and the associated *Douglas-Rachford envelope* (DRE) was introduced. Here, DRS was shown to be equivalent to a (scaled) gradient method over the DRE: when one of the two summands in the cost is quadratic, then the DRE is convex. This interpretation allowed to show a global sublinear rate for the cost of order  $O(1/k)$  for DRS, and a global linear

<sup>1</sup><http://kul-forbes.github.io/ForBES/>

rate in case the quadratic term is strongly convex. Furthermore, an optimal stepsize selection criterion was obtained, as well as an accelerated algorithm with rate  $O(1/k^2)$  based on a simple additional extrapolation step.

## 6.1 Future directions

Several directions are open for further investigate on proximal envelopes. In the case of the FBE, in this thesis we were interested in directions of superlinear convergence, and we focused on quasi-Newton formulas. Several other methods can be adapted to nonsmooth problems through proximal envelopes, including nonlinear CG methods and semismooth Newton methods. Ongoing research is investigating the properties of the FBE in the case where both summands in the problem are allowed to be nonconvex [150], and this will surely broaden the range of application of FBE-based methods. Furthermore, although the variety of problems addressed in this thesis is very comprehensive, not all problems fit this framework. One idea to extend the range of applications address is that of employing fast converging algorithms, like the ones presented in Chapters 2 to 4, to solve the inner subproblems arising in augmented Lagrangian methods.

In the case of the DRE, several aspects need to be addressed. For example, Newton-type methods can also be employed to minimize the DRE: a second-order analysis, such as the one presented in Chapter 2 for the FBE, would provide also in this case conditions for fast asymptotic convergence. Furthermore, the duality between DRS and the alternating directions method of multipliers (ADMM) should be exploited, in a similar way to what is presented in Chapter 4, to derive analogously improved versions of ADMM. Given that the latter has been successfully applied to a very wide range of problems, there is no doubt that research along these lines will have significant impact.

# References

- [1] J. Andersson, J. Åkesson, and M. Diehl. CasADi: A symbolic package for automatic differentiation and optimal control. In *Recent advances in algorithmic differentiation*, pages 297–307. Springer, 2012. [78](#), [95](#)
- [2] F. J. Aragón Artacho and M. H. Geoffroy. Characterization of metric regularity of subdifferentials. *Journal of Convex Analysis*, 15(2):365–380, 2008. [117](#)
- [3] A. Argyriou, M. Signoretto, and J. A. K. Suykens. Hybrid algorithms with applications to sparse and low rank regularization. In A. Argyriou, M. Signoretto, and J. A. K. Suykens, editors, *Regularization, Optimization, Kernels, and Support Vector Machines*, pages 53–82. Chapman & Hall/CRC, 2014. [3](#)
- [4] H. Attouch and J. Bolte. On the convergence of the proximal algorithm for nonsmooth functions involving analytic features. *Mathematical Programming*, 116(1):5–16, 2009. [21](#), [47](#)
- [5] H. Attouch, J. Bolte, P. Redont, and A. Soubeyran. Proximal alternating minimization and projection methods for nonconvex problems: An approach based on the Kurdyka-Łojasiewicz inequality. *Mathematics of Operations Research*, 35(2):438–457, 2010. [21](#), [47](#)
- [6] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems: proximal algorithms, forward-backward splitting, and regularized Gauss–Seidel methods. *Mathematical Programming*, 137(1):91–129, 2013. [7](#), [21](#), [47](#), [78](#), [79](#)
- [7] J. Barzilai and J. M. Borwein. Two-point step size gradient methods. *IMA Journal of Numerical Analysis*, 8(1):141–148, 1988. [10](#)

- [8] H. H. Bauschke, J. M. Borwein, and W. Li. Strong conical hull intersection property, bounded linear regularity, Jameson’s property (G), and error bounds in convex optimization. *Mathematical Programming*, 86(1):135–160, 1999. [137](#), [138](#)
- [9] H. H. Bauschke and P. L. Combettes. *Convex analysis and monotone operator theory in Hilbert spaces*. Springer, 2011. [6](#), [7](#), [8](#), [15](#), [16](#), [24](#), [103](#), [104](#), [126](#), [140](#), [142](#), [145](#), [159](#), [160](#)
- [10] A. Beck and M. Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM Journal on Imaging Sciences*, 2(1):183–202, 2009. [7](#), [21](#), [40](#), [100](#), [113](#)
- [11] A. Beck and M. Teboulle. A fast dual proximal gradient algorithm for convex minimization and applications. *Operations Research Letters*, 42(1):1–6, 2014. [7](#), [100](#), [104](#), [105](#), [113](#)
- [12] S. Becker and J. Fadili. A quasi-Newton proximal splitting method. In *Advances in Neural Information Processing Systems*, pages 2618–2626, 2012. [11](#), [21](#), [22](#)
- [13] A. Bemporad, A. Casavola, and E. Mosca. Nonlinear control of constrained linear systems via predictive reference management. *IEEE transactions on Automatic Control*, 42(3):340–349, 1997. [130](#)
- [14] D. S. Bernstein. *Matrix mathematics: theory, facts, and formulas with application to linear systems theory*. Princeton University Press, Woodstock, 2009. [32](#), [34](#), [121](#), [136](#)
- [15] D. P. Bertsekas. *Constrained Optimization and Lagrange Multiplier Methods*. Athena Scientific, 1996. [9](#)
- [16] D. P. Bertsekas. *Nonlinear programming*. Athena Scientific, 1999. [9](#), [23](#), [60](#), [77](#), [84](#), [86](#)
- [17] D. P. Bertsekas. *Convex optimization theory*. Athena Scientific, 2009. [13](#), [150](#), [153](#)
- [18] D. P. Bertsekas. *Convex optimization algorithms*. Athena Scientific, 2015. [5](#), [6](#), [112](#)
- [19] D. P. Bertsekas and J. N. Tsitsiklis. *Parallel and Distributed Computation: Numerical Methods*. Athena Scientific, 1997. [4](#)



- [20] J. Bolte, A. Daniilidis, and A. Lewis. The Łojasiewicz inequality for non-smooth subanalytic functions with applications to subgradient dynamical systems. *SIAM Journal on Optimization*, 17(4):1205–1223, 2007. [48](#), [50](#)
- [21] J. Bolte, S. Sabach, and M. Teboulle. Proximal alternating linearized minimization for nonconvex and nonsmooth problems. *Mathematical Programming*, 146(1-2):459–494, 2014. [21](#), [47](#), [75](#)
- [22] J. F. Bonnans, J. C. Gilbert, C. Lemaréchal, and C. A. Sagastizábal. A family of variable metric proximal methods. *Mathematical Programming*, 68(1):15–47, 1995. [9](#), [36](#)
- [23] R. I. Boţ, E. R. Csetnek, and S. C. László. An inertial forward–backward algorithm for the minimization of the sum of two nonconvex functions. *EURO Journal on Computational Optimization*, pages 1–23, 2014. [68](#)
- [24] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein. Distributed optimization and statistical learning via the alternating direction method of multipliers. *Foundations and Trends® in Machine Learning*, 3(1):1–122, 2011. [4](#), [8](#), [99](#), [140](#)
- [25] S. P. Boyd and L. Vandenberghe. *Convex optimization*. Cambridge University Press, Cambridge, UK; New York, 2004. [3](#)
- [26] C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Mathematics of Computation*, 19(92):577–593, 1965. [108](#)
- [27] R. E. Bruck. An iterative solution of a variational inequality for certain monotone operators in Hilbert space. *Bulletin of the American Mathematical Society*, 81(5):890–893, 1975. [7](#)
- [28] J. V. Burke and M. Qian. On the superlinear convergence of the variable metric proximal point algorithm using Broyden and BFGS matrix secant updating. *Mathematical Programming*, 88(1):157–181, 2000. [9](#), [36](#)
- [29] R. H. Byrd, P. Lu, J. Nocedal, and C. Zhu. A limited memory algorithm for bound constrained optimization. *SIAM Journal on Scientific Computing*, 16(5):1190–1208, 1995. [80](#)
- [30] R. H. Byrd and J. Nocedal. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *SIAM Journal on Numerical Analysis*, 26(3):727–739, 1989. [58](#), [108](#)

- [31] P. H. Calamai and J. J. Moré. Projected gradient methods for linearly constrained problems. *Mathematical Programming*, 39(1):93–116, 1987. [80](#)
- [32] E. Candès and B. Recht. Exact matrix completion via convex optimization. *Communications of the ACM*, 55(6):111–119, 2012. [3](#)
- [33] E. J. Candès and Y. Plan. Matrix completion with noise. *Proceedings of the IEEE*, 98(6):925–936, 2010. [3](#)
- [34] E. J. Candès and B. Recht. Exact matrix completion via convex optimization. *Foundations of computational mathematics*, 9(6):717–772, 2009. [3](#), [66](#)
- [35] V. Chandrasekaran, S. Sanghavi, P. A. Parrilo, and A. S. Willsky. Sparse and low-rank matrix decompositions. *IFAC Proceedings Volumes*, 42(10):1493–1498, 2009. [3](#)
- [36] X. Chen and M. Fukushima. Proximal quasi-Newton methods for nondifferentiable convex optimization. *Mathematical Programming*, 85(2):313–334, 1999. [9](#), [36](#)
- [37] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In *Fixed-point algorithms for inverse problems in science and engineering*, pages 185–212. Springer, 2011. [15](#), [20](#), [140](#)
- [38] P. L. Combettes and V. R. Wajs. Signal recovery by proximal forward-backward splitting. *Multiscale Modeling & Simulation*, 4(4):1168–1200, 2005. [15](#)
- [39] Y.-H. Dai. Convergence Properties of the BFGS Algorithm. *SIAM Journal on Optimization*, 13(3):693–701, 2002. [23](#), [38](#), [53](#)
- [40] Y. H. Dai. A perfect example for the BFGS method. *Mathematical Programming*, 138:501–530, 2013. [23](#), [38](#), [53](#)
- [41] Y.-H. Dai, W. W. Hager, K. Schittkowski, and H. Zhang. The cyclic Barzilai–Borwein method for unconstrained optimization. *IMA Journal of Numerical Analysis*, 26(3):604–627, 2005. [10](#)
- [42] Y.-H. Dai and Y. Yuan. A nonlinear conjugate gradient method with a strong global convergence property. *SIAM Journal on Optimization*, 10(1):177–182, 1999. [10](#), [61](#)
- [43] Y.-H. Dai and Y. Yuan. An efficient hybrid conjugate gradient method for unconstrained optimization. *Annals of Operations Research*, 103(1):33–47, 2001. [10](#)

- [44] D. Davis and W. Yin. Convergence rate analysis of several splitting schemes. In *Splitting Methods in Communication, Imaging, Science, and Engineering*, pages 115–163. Springer Nature, 2016. 8, 11
- [45] D. Davis and W. Yin. Faster convergence rates of relaxed peaceman-rachford and ADMM under regularity assumptions. *Mathematics of Operations Research*, 2017. Online: <http://dx.doi.org/10.1287/moor.2016.0827>. 141
- [46] W. Deng and W. Yin. On the global and linear convergence of the generalized alternating direction method of multipliers. *Journal of Scientific Computing*, 66(3):889–916, 2015. 141
- [47] J. E. Dennis and J. J. Moré. A characterization of superlinear convergence and its application to quasi-Newton methods. *Mathematics of computation*, 28(126):549–560, 1974. 22, 53, 58
- [48] J. E. Dennis and J. J. Moré. Quasi-Newton methods, motivation and theory. *SIAM review*, 19(1):46–89, 1977. 10, 52, 57
- [49] J. E. Dennis and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16. SIAM, 1996. 19
- [50] M. Diehl, H. G. Bock, H. Diedam, and P.-B. Wieber. Fast direct multiple shooting algorithms for optimal robot control. In *Fast motions in biomechanics and robotics*, pages 65–93. Springer, 2006. 77
- [51] A. Domahidi, E. Chu, and S. Boyd. ECOS: An SOCP solver for embedded systems. In *European Control Conference (ECC)*, pages 3071–3076, 2013. 131
- [52] A. L. Dontchev and R. T. Rockafellar. *Implicit functions and solution mappings*, volume 208. Springer, 2009. 17, 116, 117
- [53] J. Douglas and H. H. Rachford. On the numerical solution of heat conduction problems in two and three space variables. *Transactions of the American Mathematical Society*, 82(2):421–439, 1956. 7
- [54] D. Drusvyatskiy and A. S. Lewis. Error bounds, quadratic growth, and linear convergence of proximal methods. *ArXiv e-prints*, 2016. 116, 117, 118, 125
- [55] J. Dunn. A projected Newton method for minimization problems with nonlinear inequality constraints. *Numerische Mathematik*, 53(4):377–409, 1988. 80

- [56] J. C. Dunn and D. P. Bertsekas. Efficient dynamic programming implementations of Newton’s method for unconstrained optimal control problems. *Journal of Optimization Theory and Applications*, 63(1):23–38, 1989. [77](#), [89](#)
- [57] J. Eckstein. *Splitting methods for monotone operators with applications to parallel optimization*. Phd thesis, Massachusetts Institute of Technology, 1989. [6](#), [8](#), [140](#)
- [58] J. Eckstein and D. P. Bertsekas. On the Douglas—Rachford splitting method and the proximal point algorithm for maximal monotone operators. *Mathematical Programming*, 55(1-3):293–318, 1992. [8](#), [140](#)
- [59] F. Facchinei and J.-S. Pang. *Finite-dimensional variational inequalities and complementarity problems*, volume II. Springer, 2003. [22](#), [124](#)
- [60] M. Fazel, T. K. Pong, D. Sun, and P. Tseng. Hankel matrix rank minimization with applications to system identification and realization. *SIAM Journal on Matrix Analysis and Applications*, 34(3):946–977, 2013. [99](#)
- [61] W. Fenchel. On conjugate convex functions. *Canadian Journal of Mathematics*, 1(1):73–77, 1949. [15](#)
- [62] H. J. Ferreau, C. Kirches, A. Potschka, H. G. Bock, and M. Diehl. qpOASES: A parametric active-set algorithm for quadratic programming. *Mathematical Programming Computation*, 6(4):327–363, 2014. [131](#)
- [63] R. Fletcher. *Practical methods of optimization*. John Wiley & Sons, 1987. [60](#)
- [64] R. Fletcher. On the Barzilai-Borwein method. In *Applied Optimization*, volume 96, pages 235–256. Springer Nature, 1993. [10](#)
- [65] J. Friedman, T. Hastie, and R. Tibshirani. *The elements of statistical learning*, volume 1. Springer, 2009. [2](#), [4](#)
- [66] M. Fukushima. Equivalent differentiable optimization problems and descent methods for asymmetric variational inequality problems. *Mathematical programming*, 53(1):99–110, 1992. [22](#)
- [67] M. Fukushima and L. Qi. A globally and superlinearly convergent algorithm for nonsmooth convex minimization. *SIAM Journal on Optimization*, 6(4):1106–1120, 1996. [9](#), [36](#)
- [68] D. Gabay. Applications of the method of multipliers to variational inequalities. In M. Fortin and R. Glowinski, editors, *Augmented Lagrangian Meth-*

*ods: Applications to the Solution of Boundary-Value Problems*. North-Holland: Amsterdam, 1983. [100](#), [140](#)

- [69] E. M. Gafni and D. P. Bertsekas. Two-metric projection methods for constrained optimization. *SIAM Journal on Control and Optimization*, 22(6):936–964, 1984. [80](#)
- [70] E. Ghadimi, A. Teixeira, I. Shames, and M. Johansson. Optimal parameter selection for the alternating direction method of multipliers (ADMM): Quadratic problems. *IEEE Transactions on Automatic Control*, 60(3):644–658, 2015. [141](#)
- [71] J. C. Gilbert and J. Nocedal. Global convergence properties of conjugate gradient methods for optimization. *SIAM Journal on optimization*, 2(1):21–42, 1992. [10](#)
- [72] P. Giselsson and S. Boyd. Diagonal scaling in Douglas-Rachford splitting and ADMM. In *53rd IEEE Conference on Decision and Control*. IEEE, 2014. [8](#)
- [73] P. Giselsson and S. Boyd. Metric selection in fast dual forward–backward splitting. *Automatica*, 62:1–10, 2015. [8](#), [128](#), [130](#), [131](#)
- [74] T. Goldstein, B. O’Donoghue, S. Setzer, and R. Baraniuk. Fast alternating direction optimization methods. *SIAM Journal on Imaging Sciences*, 7(3):1588–1623, 2014. [141](#)
- [75] C. C. Gonzaga, E. W. Karas, and D. R. Rossetto. An optimal algorithm for constrained differentiable convex optimization. *SIAM Journal on Optimization*, 23(4):1939–1955, 2013. [155](#)
- [76] W. W. Hager and H. Zhang. A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM Journal on Optimization*, 16(1):170–192, 2005. [10](#)
- [77] W. W. Hager and H. Zhang. A survey of nonlinear conjugate gradient methods. *Pacific journal of Optimization*, 2(1):35–58, 2006. [10](#)
- [78] B. He and X. Yuan. On the  $O(1/n)$  convergence rate of the Douglas–Rachford alternating direction method. *SIAM Journal on Numerical Analysis*, 50(2):700–709, 2012. [141](#)
- [79] M. R. Hestenes. Multiplier and gradient methods. *Journal of optimization theory and applications*, 4(5):303–320, 1969. [111](#)

- [80] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms I: fundamentals*, volume 305. Springer Science & Business Media, 1996. [60](#), [103](#)
- [81] J.-B. Hiriart-Urruty and C. Lemaréchal. *Fundamentals of Convex Analysis*. Springer Science & Business Media, 2001. [15](#), [107](#), [135](#)
- [82] M. Hong and Z.-Q. Luo. On the linear convergence of the alternating direction method of multipliers. *Mathematical Programming*, 162(1):165–199, 2017. [141](#)
- [83] R. A. Horn and C. R. Johnson. *Matrix analysis*. Cambridge university press, 2012. [35](#), [124](#)
- [84] C.-M. Ip and J. Kyriasis. Local convergence of quasi-Newton methods for B-differentiable equations. *Mathematical Programming*, 56(1-3):71–89, 1992. [16](#), [53](#), [73](#)
- [85] K. Kurdyka. On gradients of functions definable in o-minimal structures. *Annales de l’institut Fourier*, 48(3):769–783, 1998. [21](#), [47](#)
- [86] J. Lee, Y. Sun, and M. Saunders. Proximal Newton-type methods for convex optimization. In *Advances in Neural Information Processing Systems*, pages 836–844, 2012. [11](#), [21](#), [22](#)
- [87] C. Lemaréchal and C. Sagastizábal. Practical aspects of the Moreau–Yosida regularization: Theoretical preliminaries. *SIAM Journal on Optimization*, 7(2):367–385, 1997. [32](#), [143](#)
- [88] D.-H. Li and M. Fukushima. On the global convergence of the BFGS method for nonconvex unconstrained optimization problems. *SIAM Journal on Optimization*, 11(4):1054–1064, 2001. [59](#)
- [89] W. Li. Error bounds for piecewise convex quadratic programs and applications. *SIAM Journal on Control and Optimization*, 33(5):1510–1529, 1995. [116](#)
- [90] W. Li and J. Peng. Exact penalty functions for constrained minimization problems via regularized gap function for variational inequalities. *Journal of Global Optimization*, 37(1):85–94, 2007. [22](#)
- [91] C.-J. Lin and J. J. Moré. Newton’s method for large bound-constrained optimization problems. *SIAM Journal on Optimization*, 9(4):1100–1127, 1999. [80](#)

- [92] P.-L. Lions and B. Mercier. Splitting algorithms for the sum of two nonlinear operators. *SIAM Journal on Numerical Analysis*, 16(6):964–979, 1979. 6, 7, 20, 139, 141
- [93] D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 1989. 10, 53, 59, 79, 109
- [94] T. Liu and T. K. Pong. Further properties of the forward-backward envelope with applications to difference-of-convex programming. *Computational Optimization and Applications*, 2017. Online: <http://dx.doi.org/10.1007/s10589-017-9900-2>. 102
- [95] S. Łojasiewicz. Une propriété topologique des sous-ensembles analytiques réels. *Les équations aux dérivées partielles*, pages 87–89, 1963. 21, 47
- [96] S. Łojasiewicz. Sur la géométrie semi- et sous-analytique. *Annales de l'institut Fourier*, 43(5):1575–1595, 1993. 21, 47
- [97] D. Lorenz. Constructing test instances for basis pursuit denoising. *IEEE Transactions on Signal Processing*, 61(5):1210–1214, 2013. 156
- [98] W. F. Mascarenhas. The BFGS method with exact line searches fails for non-convex objective functions. *Mathematical Programming*, 99(1):49–61, 2004. 23, 38, 53
- [99] W. F. Mascarenhas. On the divergence of line search methods. *Computational & Applied Mathematics*, 26:129 – 169, 2007. 23, 38, 53
- [100] R. Mifflin, D. Sun, and L. Qi. Quasi-Newton bundle-type methods for non-differentiable convex optimization. *SIAM Journal on Optimization*, 8(2):583–603, 1998. 9, 36
- [101] J.-J. Moreau. Proximité et dualité dans un espace Hilbertien. *Bulletin de la Société mathématique de France*, 93:273–299, 1965. 5, 8, 14, 15, 20, 21
- [102] A. Nedić and A. Ozdaglar. Approximate primal solutions and rate analysis for dual subgradient methods. *SIAM Journal on Optimization*, 19(4):1757–1780, 2009. 5
- [103] A. Nedic and A. Ozdaglar. Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, 54(1):48–61, 2009. 5

- [104] Y. Nesterov. A method of solving a convex programming problem with convergence rate  $O(1/k^2)$ . *Soviet Mathematics Doklady*, 27(2):372–376, 1983. [7](#), [21](#), [105](#), [142](#), [152](#)
- [105] Y. Nesterov. *Introductory lectures on convex optimization: A basic course*, volume 87. Springer, 2003. [105](#), [142](#), [151](#), [152](#), [153](#)
- [106] Y. Nesterov. Gradient methods for minimizing composite functions. *Mathematical Programming*, 140(1):125–161, 2013. [7](#), [21](#), [24](#), [44](#), [100](#), [105](#), [114](#)
- [107] J. Nocedal. Updating quasi-Newton matrices with limited storage. *Mathematics of computation*, 35(151):773–782, 1980. [59](#), [109](#)
- [108] J. Nocedal and S. Wright. *Numerical Optimization*. Springer, New York, 2nd edition edition, 2006. [10](#), [41](#), [57](#), [59](#), [60](#), [77](#), [109](#)
- [109] D. Noll and A. Rondepierre. *Convergence of Linesearch and Trust-Region Methods Using the Kurdyka–Łojasiewicz Inequality*, pages 593–611. Springer, New York, 2013. [43](#)
- [110] P. Ochs, Y. Chen, T. Brox, and T. Pock. iPiano: Inertial proximal algorithm for nonconvex optimization. *SIAM Journal on Imaging Sciences*, 7(2):1388–1419, 2014. [21](#), [47](#)
- [111] J.-S. Pang. Newton’s method for B-differentiable equations. *Mathematics of Operations Research*, 15(2):311–341, 1990. [16](#), [17](#)
- [112] N. Parikh and S. Boyd. Proximal algorithms. *Foundations and Trends® in Optimization*, 1(3):127–239, 2014. [5](#), [6](#), [7](#), [8](#), [99](#), [140](#)
- [113] P. Patrinos and A. Bemporad. Proximal Newton methods for convex composite optimization. In *52nd IEEE Conference on Decision and Control*, pages 2358–2363, 2013. [10](#), [21](#), [78](#), [81](#), [100](#), [110](#), [142](#), [147](#)
- [114] P. Patrinos and A. Bemporad. An accelerated dual gradient-projection algorithm for embedded linear model predictive control. *IEEE Transactions on Automatic Control*, 59(1):18–33, 2014. [4](#), [128](#), [129](#)
- [115] P. Patrinos, P. Sopasakis, and H. Sarimveis. A global piecewise smooth Newton method for fast large-scale model predictive control. *Automatica*, 47:2016–2022, 2011. [22](#)
- [116] P. Patrinos, L. Stella, and A. Bemporad. Douglas-Rachford splitting: Complexity estimates and accelerated variants. In *53rd IEEE Conference on Decision and Control (CDC)*, pages 4234–4239, 2014. [11](#), [12](#)



- [117] P. Patrinos, L. Stella, and A. Bemporad. Forward-backward truncated Newton methods for convex composite optimization. *arXiv:1402.6655*, 2014. [81](#), [98](#), [142](#)
- [118] R. A. Poliquin and R. T. Rockafellar. Amenable functions in optimization. *Nonsmooth optimization: methods and applications (Erice, 1991)*, pages 338–353, 1992. [18](#), [32](#), [87](#)
- [119] R. A. Poliquin and R. T. Rockafellar. Second-order nonsmooth analysis in nonlinear programming. *Recent advances in nonsmooth optimization*, pages 322–349, 1995. [18](#), [32](#), [87](#)
- [120] R. A. Poliquin and R. T. Rockafellar. Generalized Hessian properties of regularized nonsmooth functions. *SIAM Journal on Optimization*, 6(4):1121–1137, 1996. [18](#), [33](#), [120](#), [136](#)
- [121] M. J. D. Powell. A method for nonlinear constraints in minimization problems. In R. Fletcher, editor, *Optimization*, pages 283–298. Academic Press, New York, 1969. [111](#)
- [122] M. J. D. Powell. A hybrid method for nonlinear equations. *Numerical Methods for Nonlinear Algebraic Equations*, pages 87–144, 1970. [108](#)
- [123] M. J. D. Powell. Some global convergence properties of a variable metric algorithm for minimization without exact line searches. *Nonlinear programming*, 9:53–72, 1976. [53](#)
- [124] R. Quirynen. *Numerical Simulation Methods for Embedded Optimization*. PhD thesis, KU Leuven, 2017. [77](#)
- [125] M. Raydan. On the Barzilai and Borwein choice of steplength for the gradient method. *IMA Journal of Numerical Analysis*, 13(3):321–326, 1993. [10](#)
- [126] S. Richter, C. N. Jones, and M. Morari. Certification aspects of the fast gradient method for solving the dual of parametric convex programs. *Mathematical Methods of Operations Research*, 77(3):305–321, 2013. [131](#)
- [127] R. T. Rockafellar. A dual approach to solving nonlinear programming problems by unconstrained optimization. *Mathematical Programming*, 5(1):354–373, 1973. [9](#), [111](#)
- [128] R. T. Rockafellar. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of operations research*, 1(2):97–116, 1976. [9](#), [111](#)

- [129] R. T. Rockafellar. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization*, 14(5):877–898, 1976. [6](#), [36](#), [139](#), [140](#)
- [130] R. T. Rockafellar. First- and second-order epi-differentiability in nonlinear programming. *Transactions of the American Mathematical Society*, 307(1):75–108, 1988. [18](#), [32](#), [87](#)
- [131] R. T. Rockafellar. Second-order optimality conditions in nonlinear programming obtained by way of epi-derivatives. *Mathematics of Operations Research*, 14(3):462–484, 1989. [18](#), [32](#), [87](#)
- [132] R. T. Rockafellar. *Convex Analysis*. Princeton university press, 1997. [13](#), [15](#), [103](#), [137](#)
- [133] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317. Springer, 2011. [14](#), [15](#), [16](#), [18](#), [27](#), [32](#), [34](#), [35](#), [57](#), [82](#), [103](#), [107](#), [116](#), [118](#), [120](#), [121](#), [122](#), [123](#), [135](#), [136](#)
- [134] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, 1992. [4](#)
- [135] N. Sagara and M. Fukushima. A trust region method for nonsmooth convex optimization. *Management*, 1(2):171–180, 2005. [9](#), [36](#)
- [136] A. K. Sampathirao, P. Sopasakis, A. Bemporad, and P. Patrinos. Proximal limited-memory quasi-Newton methods for scenario-based stochastic optimal control. *To appear in Proceedings of the 20th IFAC World Congress*, 2017. [102](#)
- [137] K. Scheinberg and X. Tang. Practical inexact proximal quasi-newton method with global complexity analysis. *Mathematical Programming*, 160(1-2):495–529, 2016. [11](#), [21](#), [22](#)
- [138] M. Schmidt, D. Kim, and S. Sra. Projected Newton-type methods in machine learning. In S. Sra, S. Nowozin, and S. J. Wright, editors, *Optimization for Machine Learning*, pages 305–329. MIT Press, 2011. [11](#)
- [139] F. Schöpfer. Linear convergence of descent methods for the unconstrained minimization of restricted strongly convex functions. *SIAM Journal on Optimization*, 26(3):1883–1911, 2016. [116](#)

- [140] N. Z. Shor. *Minimization Methods for Non-Differentiable Functions*. Springer, Berlin, 1985. [5](#)
- [141] M. Signoretto, Q. T. Dinh, L. D. Lathauwer, and J. A. K. Suykens. Learning with tensors: a framework based on convex optimization and spectral regularization. *Machine Learning*, 94(3):303–351, 2013. [3](#)
- [142] P. Sotasakis, N. Freris, and P. Patrinos. Accelerated reconstruction of a compressively sampled data stream. *arXiv:1605.02341*, 2016. [156](#)
- [143] W. Squire and G. Trapp. Using complex variables to estimate derivatives of real functions. *SIAM Review*, 40(1):110–112, 1998. [41](#)
- [144] G. Stathopoulos, H. Shukla, A. Szucs, Y. Pu, and C. N. Jones. Operator splitting methods in control. *Foundations and Trends® in Systems and Control*, 3(3):249–362, 2016. [4](#), [99](#), [100](#), [128](#)
- [145] L. Stella, A. Themelis, and P. Patrinos. Forward-backward quasi-Newton methods for nonsmooth optimization problems. *Computational Optimization and Applications*, 67(3):443–487, 2017. [10](#), [11](#), [78](#), [81](#), [100](#), [102](#), [109](#), [114](#)
- [146] L. Stella, A. Themelis, and P. Patrinos. Newton-type alternating minimization algorithm for convex optimization. Submitted to IEEE Transactions on Automatic Control (TAC), 2017. [10](#), [12](#)
- [147] L. Stella, A. Themelis, P. Sotasakis, and P. Patrinos. A simple and efficient algorithm for nonlinear model predictive control. Submitted to the 56th IEEE Conference on Decision and Control (CDC), 2017. [10](#), [12](#)
- [148] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization methods and software*, 11(1-4):625–653, 1999. [131](#)
- [149] A. Themelis and P. Patrinos. SuperMann: a superlinearly convergent algorithm for finding fixed points of nonexpansive operators. *arXiv:1609.06955*, 2016. [126](#)
- [150] A. Themelis, L. Stella, and P. Patrinos. Forward-backward envelope for the sum of two nonconvex functions: Further properties and nonmonotone line-search algorithms. *arXiv:1606.06256*, 2016. [78](#), [81](#), [83](#), [100](#), [102](#), [164](#)
- [151] R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996. [2](#)

- [152] K.-C. Toh, M. J. Todd, and R. H. Tütüncü. SDPT3 – a MATLAB software package for semidefinite programming, version 1.3. *Optimization methods and software*, 11(1-4):545–581, 1999. [131](#)
- [153] K.-C. Toh and S. Yun. An accelerated proximal gradient algorithm for nuclear norm regularized linear least squares problems. *Pacific Journal of Optimization*, 6(3):615–640, 2010. [67](#)
- [154] P. Tseng. Applications of a splitting algorithm to decomposition in convex programming and variational inequalities. *SIAM Journal on Control and Optimization*, 29(1):119–138, 1991. [7](#), [99](#), [100](#), [105](#)
- [155] P. Tseng. On accelerated proximal gradient methods for convex-concave optimization. Technical report, Department of Mathematics, University of Washington, 2008. [7](#), [21](#)
- [156] M. Vukov, A. Domahidi, H. J. Ferreau, M. Morari, and M. Diehl. Auto-generated algorithms for nonlinear model predictive control on long and on short horizons. In *52nd IEEE Conference on Decision and Control*, pages 5113–5118. IEEE, 2013. [92](#)
- [157] Z. Wen, W. Yin, D. Goldfarb, and Y. Zhang. A fast algorithm for sparse reconstruction based on shrinkage, subspace optimization, and continuation. *SIAM Journal on Scientific Computing*, 32(4):1832–1857, 2010. [61](#)
- [158] L. Wirsching, H. G. Bock, and M. Diehl. Fast NMPC of a chain of masses connected by springs. In *IEEE International Conference on Control Applications*, pages 591–596, 2006. [92](#), [93](#)
- [159] J. Wright, A. Ganesh, S. Rao, Y. Peng, and Y. Ma. Robust principal component analysis: Exact recovery of corrupted low-rank matrices via convex optimization. In Y. Bengio, D. Schuurmans, J. D. Lafferty, C. K. I. Williams, and A. Culotta, editors, *Advances in Neural Information Processing Systems* 22, pages 2080–2088. Curran Associates, Inc., 2009. [3](#)
- [160] S. J. Wright. Solution of discrete-time optimal control problems on parallel computers. *Parallel Computing*, 16(2-3):221–237, 1990. [77](#)
- [161] S. J. Wright, R. D. Nowak, and M. A. Figueiredo. Sparse reconstruction by separable approximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493, 2009. [61](#)

- [162] N. Yamashita, K. Taji, and M. Fukushima. Unconstrained optimization reformulations of variational inequality problems. *Journal of Optimization Theory and Applications*, 92(3):439–456, 1997. [22](#)
- [163] Z. Zhou and A. M.-C. So. A unified approach to error bounds for structured convex optimization problems. *arXiv:1512.03518*, 2015. [117](#)
- [164] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005. [2](#)





Unless otherwise expressly stated, all original material of whatever nature created by Lorenzo Stella and included in this thesis, is licensed under a [Creative Commons Attribution Noncommercial Share Alike 2.5 Italy License](https://creativecommons.org/licenses/by-nc-sa/2.5/it/).

Check [creativecommons.org/licenses/by-nc-sa/2.5/it/](https://creativecommons.org/licenses/by-nc-sa/2.5/it/) for the legal code of the full license.

[Ask the author](#) about other uses.